



TUGAS AKHIR - SS141501

# PEMODELAN FAKTOR-FAKTOR YANG MEMENGARUHI INDEKS PEMBANGUNAN MANUSIA MENGGUNAKAN REGRESI NONPARAMETRIK *SPLINE* DI JAWA TENGAH

NI PUTU DERA YANTHI  
NRP 1312 100 040

Dosen Pembimbing  
Prof. Dr. I Nyoman Budiantara, M.Si.

PROGRAM STUDI S1  
JURUSAN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2016



FINAL PROJECT - SS141501

# **MODELING FACTORS THAT INFLUENCE HUMAN DEVELOPMENT INDEX USING NONPARAMETRIC SPLINE REGRESSION AT CENTRAL JAVA**

NI PUTU DERA YANTHI  
NRP 1312 100 040

Supervisor  
Prof. Dr. I Nyoman Budiantara, M. Si.

UNDERGRADUATE PROGRAMME  
DEPARTMENT OF STATISTICS  
FACULTY OF MATHEMATICS AND NATURAL SCIENCE  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2016

## LEMBAR PENGESAHAN

# PEMODELAN FAKTOR-FAKTOR YANG MEMENGARUHI INDEKS PEMBANGUNAN MANUSIA MENGGUNAKAN REGRESI NONPARAMETRIK *SPLINE* DI JAWA TENGAH

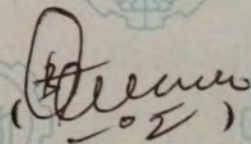
## TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Sains  
pada  
Program Studi S-1 Jurusan Statistika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Teknologi Sepuluh Nopember


Oleh :  
**NI PUTU DERA YANTHI**  
NRP 1312 100 040

Disetujui oleh Pembimbing Tugas Akhir :

Prof. Dr. I Nyoman Budiantara, M.Si  
NIP. 19650603 198903 1 003



Mengetahui  
Ketua Jurusan Statistika FMIPA-ITS



**Dr. Suhartono**

NIP. 19710929 199512 1 001

SURABAYA, JULI 2016

**PEMODELAN FAKTOR-FAKTOR YANG  
MEMENGARUHI INDEKS PEMBANGUNAN MANUSIA  
MENGUNAKAN PENDEKATAN REGRESI  
NONPARAMETRIK *SPLINE* DI JAWA TENGAH**

**Nama** : Ni Putu Dera Yanthi  
**NRP** : 1312100040  
**Jurusan** : Statistika  
**Dosen Pembimbing** : Prof. Dr. I Nyoman Budiantara, M. Si.

**Abstrak**

Manusia merupakan salah satu modal dasar yang dimiliki oleh suatu bangsa. Pada tahun 1990, Badan Perserikatan Bangsa-Bangsa (PBB) menetapkan suatu ukuran standar pembangunan manusia yaitu Indeks Pembangunan Manusia (IPM) atau *Human Development Index* (HDI). Pada tahun 2013, Jawa Tengah memiliki nilai IPM terendah setelah Jawa Timur di Pulau Jawa dengan nilai IPM sebesar 68,02. Hal ini menunjukkan bahwa perkembangan pembangunan di provinsi Jawa Tengah masih tertinggal dari provinsi-provinsi yang terdapat di Pulau Jawa. Tindakan yang dapat dilakukan pemerintah untuk meningkatkan nilai IPM adalah dengan mengetahui faktor-faktor yang memengaruhi IPM di Jawa Tengah. Metode statistika yang digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor adalah analisis regresi. Pada penelitian ini, pola data IPM dan faktor-faktor yang memengaruhi IPM di provinsi Jawa Tengah dilihat dari *scatterplot* memiliki pola data yang tidak diketahui bentuk polanya sehingga metode yang dapat digunakan adalah regresi nonparametrik *spline*. Model terbaik didapatkan dari titik knot optimal berdasarkan nilai *Generalized Cross Validation* (GCV) terkecil. Berdasarkan penelitian ini, model regresi nonparametrik *spline* terbaik adalah dengan menggunakan kombinasi knot (3,3,2,1,2) dan lima variabel signifikan yaitu tingkat partisipasi angkatan kerja, rasio sekolah-siswa, kepadatan penduduk, angka kesakitan, dan PDRB/1juta setiap kabupaten/kota di Jawa Tengah. Model tersebut memiliki  $R^2$  sebesar 93,14% dan MSE sebesar 6,45564.

Kata kunci: GCV, Indeks Pembangunan Manusia, Regresi Nonparametrik, *Spline*, Titik knot

# MODELING FACTORS THAT INFLUENCE HUMAN DEVELOPMENT INDEX USING NONPARAMETRIC SPLINE REGRESSION AT CENTRAL JAVA

**Name** : Ni Putu Dera Yanthi  
**NRP** : 1312100040  
**Department** : Statistics  
**Supervisor** : Prof. Dr. I Nyoman Budiantara, M. Si.

## *Abstract*

*Human is an asset that belongs to a nation. In 1990, United Nations considered one standard measure for Human Development Index (HDI). In 2013, Central Java have the lowest HDI after East Java in Java Island with 68.02. This shows that development progress in Central Java are far behind other provinces at Java Island. An action from government to increase HDI is by identifying factors that influence HDI in Central Java. Statistical method used to define relation between response variable and dependent variable is regression analysis. In this research, HDI pattern and factors that influence HDI in Central Java shown from scatterplot that do not have a certain pattern and therefore nonparametric spline regression conducted. Best model from optimum knots are based on the lowest Generalized Cross Validation (GCV) value. From this research, the best nonparametric spline regression model with knot combination of (3,3,2,1,2) and five significant variables are labor force participation rates, the ratio of students, population density, morbidity, and regional gross income for each million for district/city in Central Java. The model have the  $R^2$  of 93.14% and MSE of 6.45564.*

*Keywords: GCV, Human Development Index, Knots, Nonparametric Regression, Spline*

# DAFTAR ISI

halaman

<b>HALAMAN JUDUL</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	iii
<b>ABSTRAK</b> .....	v
<b>ABSTRACT</b> .....	vii
<b>KATA PENGANTAR</b> .....	ix
<b>DAFTAR ISI</b> .....	xi
<b>DAFTAR TABEL</b> .....	xiii
<b>DAFTAR GAMBAR</b> .....	xv
<b>DAFTAR LAMPIRAN</b> .....	xvii
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	4
1.5 Batasan Masalah .....	5
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Statistika Deskriptif.....	7
2.2 Analisis Regresi .....	7
2.3 Regresi Nonparametrik Spline .....	8
2.4 Pemilihan Titik Knot Optimal.....	9
2.5 Kriteria Pemilihan Model Terbaik .....	10
2.6 Estimasi Parameter.....	11
2.7 Pengujian Parameter Model .....	11
a. Pengujian Secara Serentak .....	11
b. Pengujian Secara Individu.....	12
2.8 Pengujian Asumsi Residual.....	13
a. Uji Asumsi Identik .....	13
b. Uji Asumsi Independen .....	14
c. Uji Asumsi Distribusi Normal.....	15
2.8 Indeks Pembangunan Manusia .....	16

### **BAB III METODOLOGI PENELITIAN**

3.1 Sumber Data.....	19
3.2 Variabel Penelitian.....	19
3.3 Langkah Analisis.....	21
3.4 Diagram Alir .....	21

### **BAB IV ANALISIS DAN PEMBAHASAN**

4.1 Karakteristik Indeks Pembangunan Manusia di Jawa Tengah.....	23
4.2 Analisis Faktor-faktor yang Diduga Memengaruhi Indeks Pembangunan Manusia.....	25
4.3 Regresi Nonparametrik .....	29
4.3.1 Pemilihan Titik Knot Otimum.....	29
1. Pemilihan Titik Knot dengan Satu Titik Knot.....	30
2. Pemilihan Titik Knot dengan Dua Titik Knot.....	30
3. Pemilihan Titik Knot dengan Tiga Titik Knot .....	32
4. Pemilihan Titik Knot dengan Kombinasi Knot .....	33
4.3.2 Pemilihan Model Terbaik.....	34
4.3.3 Pengujian Signifikansi Parameter Model Regresi Nonparametrik Spline.....	35
1. Uji Serentak.....	36
2. Uji Individu .....	36
4.3.4 Pengujian Asumsi Residual.....	37
1. Asumsi Identik .....	37
2. Asumsi Independen .....	38
3. Asumsi Distribusi Normal.....	39
4.4 Kriteria Pemilihan Model Terbaik .....	40
4.5 Pemodelan Regresi Nonparametrik <i>Spline</i> .....	40

### **BAB V KESIMPULAN DAN SARAN**

5.1 Kesimpulan .....	43
5.2 Saran.....	44

### **DAFTAR PUSTAKA .....**

### **LAMPIRAN .....**

### **BIODATA PENULIS .....**

## DAFTAR TABEL

halaman

Tabel 2.1	Analisis Ragam (ANOVA) .....	12
Tabel 2.2	Nilai Maksimum dan Minimum dari setiap Komponen IPM.....	17
Tabel 3.1	Variabel Penelitian .....	19
Tabel 4.1	Statistika Deskriptif Variabel Penelitian.....	27
Tabel 4.2	Nilai GCV dengan Satu Titik Knot.....	30
Tabel 4.3	Nilai GCV dengan Dua Titik Knot .....	31
Tabel 4.4	Nilai GCV dengan Tiga Titik Knot .....	33
Tabel 4.5	Nilai GCV dengan Kombinasi Titik Knot .....	34
Tabel 4.6	Perbandingan Nilai GCV Minimum .....	35
Tabel 4.7	Hasil Uji Serentak .....	36
Tabel 4.8	Hasil Uji Individu .....	37
Tabel 4.9	Hasil Uji Glejser .....	38



## DAFTAR GAMBAR

	halaman
Gambar 3.1 Diagram Alir .....	22
Gambar 4.1 Diagram Indeks Pembangunan Manusia di Kabupaten/Kota Provinsi Jawa Tengah .....	24
Gambar 4.2 Pola antara IPM dengan TPAK.....	25
Gambar 4.3 Pola antara IPM dengan Rasio Sekolah-Siswa	25
Gambar 4.4 Pola antara IPM dengan Kepadatan Penduduk	26
Gambar 4.5 Pola antara IPM dengan Angka Kesakitan.....	26
Gambar 4.6 Pola antara IPM dengan PDRB/1juta.....	27
Gambar 4.7 Plot ACF Residual .....	39
Gambar 4.8 Hasil Uji <i>Kolmogorv-Smirnov</i> .....	39
Gambar 4.9 Nilai $\hat{Y}_i$ dan $Y_i$ .....	41

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Manusia merupakan salah satu kekayaan yang dimiliki oleh suatu bangsa, sehingga kesejahteraan manusia harus diperhitungkan. Pada tahun 1990, *United Nations Development Programme* (UNDP) memperkenalkan konsep *Human Development* atau pembangunan manusia sebagai paradigma baru model pembangunan. Untuk dapat ikut berpartisipasi dalam proses pembangunan, tentunya dibutuhkan masyarakat Indonesia yang tidak hanya unggul dari segi kuantitas, tetapi juga unggul dari segi kualitas (BPS, 2011). Berbagai ukuran pembangunan manusia dibuat namun tidak semuanya dapat digunakan sebagai ukuran standar yang dapat dibandingkan antarwilayah dan antarnegara. Oleh karena itu, pada tahun 1990 Badan Perserikatan Bangsa-Bangsa (PBB) menetapkan suatu ukuran standar pembangunan manusia yaitu Indeks Pembangunan Manusia (IPM) atau *Human Development Index* (HDI) untuk melihat sejauh mana keberhasilan pembangunan dapat mewujudkan masyarakat yang makmur dan sejahtera. IPM mengukur pencapaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup. Sebagai ukuran hidup, IPM dibangun melalui pendekatan tiga dimensi dasar. Dimensi tersebut ialah kesehatan dengan indikator angka harapan hidup, pengetahuan dengan indikator angka melek huruf dan rata-rata lama sekolah, dan yang terakhir ialah dimensi hidup layak dengan indikator kemampuan daya beli (BPS, 2007).

IPM merupakan relasi antara manusia dengan pembangunan disekitarnya. Sehingga semakin padat penduduk dalam suatu wilayah maka angka IPM harus diperhitungkan. Di Indonesia Pulau Jawa merupakan pulau dengan penduduk yang paling padat, salah satunya adalah Jawa Tengah. Pada tahun 2013 salah satu provinsi yang berada di Pulau Jawa yaitu Jawa Tengah memiliki nilai IPM sebesar 68,02. Angka IPM di Jawa Tengah masih

dibawah dari provinsi-provinsi besar seperti Jakarta, Jawa Barat, Banten, dan Yogyakarta. Hal ini menunjukkan bahwa perkembangan pembangunan di Provinsi Jawa Tengah masih tertinggal jauh dari provinsi-provinsi yang terdapat di Pulau Jawa selain Jawa Timur. Oleh karena itu perlu diketahui faktor-faktor apa sajakah yang memengaruhi Indeks Pembangunan Manusia di Jawa Tengah sehingga pemerintah provinsi setempat bisa lebih memerhatikan dan mengupayakan program-program pembangunan manusia guna meningkatkan angka Indeks Pembangunan Manusia sebagai usaha perbaikan kesejahteraan manusia.

Faktor-faktor pengaruh tersebut dapat diketahui dengan menggunakan metode pemodelan. Salah satu metode yang dapat digunakan untuk menyelesaikan masalah pemodelan adalah analisis regresi. Analisis regresi merupakan suatu metode statistika yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan prediktor (Drapper & Smith, 1992). Pendekatan dalam metode analisis regresi ada tiga yaitu pendekatan parametrik, pendekatan nonparametrik dan pendekatan semiparametrik (Budiantara, 2005). Regresi parametrik digunakan jika kurva regresi mengikuti pola tertentu atau membentuk pola data yang jelas seperti linier, kuadratik atau kubik. Namun pada kenyataannya, antara variabel respon dengan prediktor tidak selalu memiliki pola hubungan yang tertentu. Metode yang dapat digunakan untuk mengetahui hubungan antara variabel respon dengan prediktor jika kurva regresinya tidak diketahui bentuk polanya adalah regresi nonparametrik.

Pada regresi nonparametrik, data dibiarkan mencari sendiri bentuk estimasinya sehingga subjektivitas peneliti dapat diminimalisir. Dengan kata lain, regresi nonparametrik memiliki fleksibilitas yang tinggi (Eubank, 1988). Kelebihan regresi nonparametrik *spline* ialah bisa memodelkan data yang perilakunya berubah atau pola data yang berubah-ubah pada selang tertentu yang tak mampu dimodelkan oleh parametrik. Setiap data

memiliki sebuah pola, terdapat pola yang bisa diidentifikasi dan terdapat pola yang tidak bisa diidentifikasi. Pola data yang bisa diidentifikasi ialah pola data dengan bentuk yang jelas, sedangkan pola yang tidak bisa diidentifikasi ialah pola yang tidak diketahui bentuk polanya. Pola data yang tidak diketahui bentuk polanya bisa di jelaskan dengan interval-interval yang berada pada titik x. Setiap interval memiliki perubahan bentuk pola, sehingga dalam pola data yang tidak bisa diidentifikasi terdapat perubahan bentuk pola bila terdapat pada selang tertentu.

IPM mempunyai tiga dimensi yaitu pengetahuan, kesehatan dan standar hidup layak. Sehingga variabel yang digunakan ialah yang berhubungan didalam dimensi tersebut namun tidak termasuk pada perhitungan IPM. Pada dimensi kesehatan menggunakan variabel angka kesakitan, dimensi pengetahuan menggunakan variabel rasio sekolah-siswa dan pada dimensi standar hidup layak menggunakan variabel tingkat partisipasi angkatan kerja, kepadatan penduduk, dan PDRB/1juta. Pada penelitian ini, pola data IPM dan faktor-faktor yang memengaruhi IPM di provinsi Jawa Tengah jika dilihat dari *scatterplot* memiliki pola data yang tidak diketahui bentuknya sehingga metode analisis yang digunakan adalah regresi nonparametrik *spline*.

Penelitian sebelumnya tentang IPM sudah banyak dilakukan, beberapa diantaranya yaitu penelitian yang dilakukan oleh Melliana (2013) yang menghasilkan faktor-faktor yang berpengaruh terhadap IPM di provinsi Jawa Timur menggunakan regresi panel yaitu angka partisipasi sekolah SMP, rasio guru-siswa SMP, rasio sekolah-murid SMP, jumlah sarana kesehatan, persentase rumah tangga dengan akses air bersih, tingkat partisipasi angkatan kerja, PDRB perkapita, dan kepadatan penduduk. Penelitian Awal (2014) yang menghasilkan faktor-faktor yang berpengaruh terhadap IPM di provinsi papua yaitu rasio sekolah-guru pada sekolah menengah pertama, persentase pusat kesehatan masyarakat dan sejenisnya, persentase tenaga kesehatan masyarakat dan sejenisnya, laju pertumbuhan PDRB dan

persentase penduduk miskin menggunakan metode nonparametrik. Penelitian oleh Retno (2014) yang menghasilkan faktor-faktor yang berpengaruh terhadap IPM di Jawa Timur yaitu angka kematian bayi, pertumbuhan ekonomi, tingkat pengangguran, terbuka, tingkat partisipasi angkatan kerja menggunakan metode regresi semiparametrik. Hingga saat ini belum ada yang melakukan penelitian tentang pemodelan IPM di Jawa Tengah dengan pendekatan regresi nonparametrik *spline*.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah dipaparkan mengenai indeks pembangunan manusia di Jawa Tengah, maka permasalahan yang akan dibahas pada penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik IPM di provinsi Jawa Tengah?
2. Bagaimana pemodelan IPM provinsi Jawa Tengah dengan pendekatan regresi nonparametrik *spline*?

## **1.3 Tujuan**

Berdasarkan rumusan masalah, tujuan yang ingin dicapai pada penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik IPM di provinsi Jawa Tengah.
2. Memodelkan faktor – faktor yang memengaruhi IPM provinsi Jawa Tengah dengan pendekatan regresi nonparametrik *spline*.

## **1.4 Manfaat Penelitian**

Manfaat yang diharapkan dapat diperoleh dari hasil penelitian yang dilakukan ini diantaranya sebagai berikut.

1. Memberikan informasi kepada pemerintah Jawa Tengah mengenai faktor-faktor yang memengaruhi Indeks Pembangunan Manusia sehingga nantinya dapat dijadikan pertimbangan dalam pelaksanaan program-program pemerintah dan sebagai bahan pertimbangan dalam menentukan kebijakan.

2. Mampu menerapkan ilmu yang sejalan dengan bidang Statistika yaitu metode regresi nonparametrik *spline* dalam bentuk nyata khususnya pada bidang sosial pemerintahan.
3. Menjadi bahan masukan atau acuan dalam penelitian selanjutnya yang akan dilakukan.

### **1.5 Batasan Masalah**

Dalam penelitian ini perlu diberikan batasan permasalahan agar penelitian yang akan dikerjakan lebih fokus dan sesuai dengan rentang waktu yang direncanakan. Batasan masalah dalam penelitian ini adalah sebagai berikut.

1. Data yang digunakan adalah data sekunder dari Badan Pusat Statistik provinsi Jawa Tengah. Data terdiri dari 35 kabupaten/kota di provinsi Jawa Tengah.
2. Data yang digunakan adalah data pada tahun 2013.
3. Pemilihan titik knot optimal menggunakan metode GCV (*Generalized Cross Validation*).
4. Fungsi *spline* yang digunakan adalah *spline* linear.

(Halaman ini sengaja dikosongkan)

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Statistika Deskriptif**

Statistika deskriptif adalah statistika yang digunakan untuk menganalisis data dengan cara mendeskripsikan atau menggambarkan data yang telah terkumpul sebagaimana adanya tanpa bermaksud membuat kesimpulan yang berlaku untuk umum atau generalisasi (Sugiyono, 2010). Pada penelitian ini analisis deskriptif yang digunakan adalah *scatterplot*. *Scatterplot* adalah plot dua dimensi yang menggambarkan hubungan antara dua variabel. Berdasarkan *scatterplot* dapat diketahui apakah dua variabel memiliki hubungan positif, negatif atau tidak ada hubungan. Jika plot lebih cenderung keatas maka data memiliki hubungan positif, jika plot lebih cenderung kebawah maka data memiliki hubungan negatif, dan jika plot menyebar secara acak tidak condong keatas maupun kebawah maka data tidak memiliki hubungan yang linier.

#### **2.2 Analisis Regresi**

Analisis Statistik yang dapat digunakan untuk menyelesaikan masalah kausalitas atau sebab akibat adalah analisis regresi (Kutner, Nachtsheim, & Neter, 2004). Regresi pertama-tama dipergunakan sebagai konsep statistik pada tahun 1877 oleh Sir Francis Galton yang melakukan studi tentang kecenderungan tinggi badan anak dengan orang tuanya. Hasil studi tersebut merupakan suatu kesimpulan bahwa kecenderungan tinggi badan anak yang lahir terhadap orangtuanya adalah menurun (*regress*) mengarah pada tinggi badan rata-rata penduduk. Istilah regresi pada mulanya bertujuan untuk membuat perkiraan nilai satu variabel (tinggi badan anak) terhadap satu variabel yang lain (tinggi badan orang tua). Selanjutnya berkembang menjadi alat untuk membuat perkiraan nilai suatu variabel dengan menggunakan beberapa variabel lain yang berhubungan dengan



variabel tersebut. Sehingga dalam ilmu Statistika, teknik yang umum digunakan untuk menganalisis hubungan antara dua atau lebih variabel adalah analisis regresi.

Analisis regresi merupakan sebuah metode Statistika yang memberikan penjelasan tentang pola hubungan (model) antara dua variabel atau lebih (Drapper & Smith, 1992). Setiap variabel mempunyai pola hubungan yang berbeda-beda yaitu pola linier, kuadratik dan kubik. Namun tidak semua kasus mempunyai pola tersebut atau tidak memiliki pola khusus, maka terdapat tiga model pendekatan regresi yaitu pendekatan regresi parametrik, pendekatan regresi nonparametrik dan pendekatan regresi semiparametrik.

### 2.3 Regresi Nonparametrik *Spline*

Pendekatan model regresi nonparametrik digunakan jika bentuk kurva regresi tidak memenuhi asumsi-asumsi klasik seperti pada regresi parametrik (bentuk kurva regresi diketahui) atau informasi tentang bentuk pola data dimasa lalu tidak lengkap (Eubank, 1988). Kurva regresi hanya diasumsikan halus (*smooth*) dalam artian kurva termuat dalam suatu ruang fungsi tertentu. Bentuk model regresi nonparametrik secara umum disajikan sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.1)$$

dengan  $y_i$  adalah variabel respon, fungsi  $f$  yang tidak diketahui bentuk kurva regresinya dengan  $x_i$  sebagai variabel prediktor dan  $\varepsilon_i$  adalah *error* random yang diasumsikan berdistribusi  $N(0, \sigma^2)$ .

Melakukan estimasi terhadap fungsi  $f(x_i)$  dalam regresi nonparametrik dapat dilakukan dengan berbagai pendekatan, salah satunya adalah *spline* (Wahba, 1990; Budiantara, 2009). Estimator *spline* merupakan metode yang paling banyak mendapat perhatian dari peneliti dalam beberapa tahun terakhir. Misalkan pada persamaan 2.6 kurva regresi  $f(x_i)$  dihipotesiskan dengan fungsi *spline*

berorde  $p$  dengan titik knot  $K_1, K_2, \dots, K_r$  yang diberikan oleh persamaan sebagai berikut.

$$f(x_i) = \sum_{j=0}^p \gamma_j x_i^j + \sum_{k=1}^r \gamma_{p+k} (x_i - K_k)_+^p \quad (2.2).$$

Dengan orde  $p$  ialah derajat polynomial. Dalam penelitian ini menggunakan derajat polinomial 1 dimana nilai derajat 1 merupakan kurva linier.  $K_k$  merupakan nilai dari titik knot atau nilai titik yang mengalami perubahan pola. Apabila persamaan 2.7 disubstitusikan kedalam persamaan 2.6 maka diperoleh persamaan regresi nonparametrik *spline* sebagai berikut.

$$\begin{aligned} y_i &= \sum_{j=0}^p \gamma_j x_i^j + \sum_{k=1}^r \gamma_{p+k} (x_i - K_k)_+^p + \varepsilon_i, \\ i &= 1, 2, \dots, n \end{aligned} \quad (2.3).$$

Fungsi  $(x_i - K_k)_+^p$  merupakan fungsi potongan yang diberikan oleh

$$(x_i - K_k)_+^p = \begin{cases} (x_i - K_k)^p, & x_i \geq K_k \\ 0, & x_i < K_k \end{cases} \quad (2.4).$$

## 2.4 Pemilihan Titik Knot Optimal

Banyak metode yang dapat digunakan untuk memilih titik knot optimal dalam estimator *spline*. Wahba (1990) memberikan suatu metode untuk memilih titik knot yang optimal yaitu dengan metode GCV (*Generalized Cross Validation*). Secara teoritis metode GCV mempunyai sifat optimal asimtotik yang tidak dimiliki oleh metode lainnya. Kelebihan ini menjadikan metode GCV sangat terkenal dalam regresi nonparametrik dan sering digeneralisasikan dan disesuaikan bentuk formulanya oleh para peneliti dalam estimator *spline* lain untuk memilih titik knot yang optimal. Budiantara (2000) menggeneralisasikan metode GCV dari Wahba (1990) untuk estimator *spline* terbobot dan memperlihatkan sifat optimal asimtotik masih berlaku untuk estimator yang terboboti.

Menurut Budiantara (2006) pemilihan titik-titik knot optimal dengan memilih fungsi yang memiliki nilai GCV paling minimum. Metode GCV secara umum dapat dituliskan sebagai berikut.

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{(n^{-1} \text{trace}[I - A(K_1, K_2, \dots, K_r)])^2} \quad (2.5)$$

dengan  $A$  merupakan matriks,  $A(K_1, K_2, \dots, K_r) = X(X'X)^{-1}X'$  dan rumus  $MSE$  didapatkan dengan formula sebagai berikut.

$$MSE(K_1, K_2, \dots, K_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 \quad (2.6).$$

Model *spline* dengan nilai GCV paling minimum dikatakan model *spline* yang terbaik.

## 2.5 Kriteria Pemilihan Model Terbaik

Beberapa kriteria yang digunakan untuk menentukan model regresi terbaik salah satunya ialah menggunakan koefisien determinasi ( $R^2$ ) dan  $MSE$ . Koefisien determinasi adalah nilai dari proporsi keragaman total disekitar nilai tengah  $\bar{y}$  yang dapat dijelaskan oleh model regresi (Drapper & Smith, 1992). Semakin tinggi nilai  $R^2$  yang dihasilkan suatu model, maka semakin baik pula variabel-variabel prediktor dalam model tersebut dalam menjelaskan variabilitas variabel respon. Berikut ini merupakan formula untuk menghitung  $R^2$ .

$$R^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.7).$$

Selain dengan  $R^2$ , pemilihan model terbaik juga dapat menggunakan  $MSE$ .  $MSE$  adalah rata-rata dari kesalahan peramalan dikuadratkan (Subagyo, 1986). Berikut ini merupakan formula untuk menghitung nilai  $MSE$ .

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (2.8)$$

Nilai *MSE* didapatkan dengan menggunakan metode *k-fold*, *k-fold* merupakan salah satu metode *Cross Validation*. *K-fold* dilakukan untuk membagi data menjadi data training dan data testing, dengan mengulang sebanyak *k* untuk membagi sebuah himpunan secara acak menjadi *k* subser yang saling bebas.

## 2.6 Estimasi Parameter

Estimasi merupakan proses yang menggunakan sampel statistic untuk menduga atau menaksir hubungan parameter populasi (Hasan, 2002). Salah satu cara untuk mencari estimasi parameter ialah menggunakan metode OLS (*Ordinary Least Square*). Metode OLS merupakan suatu metode untuk mendapatkan estimator parameter model dengan meminimumkan jumlah kuadrat error. Berikut ini merupakan formula dari estimasi parameter dengan menggunakan metode OLS.

$$\hat{y} = (X'X)^{-1}X'y \quad (2.9)$$

## 2.7 Pengujian Parameter Model

Pengujian ini bertujuan untuk mengetahui pengaruh apakah variabel prediktor berpengaruh secara signifikan pada model. Terdapat dua tahap pengujian parameter secara serentak dan pengujian secara individu (parsial).

### a. Pengujian Secara Serentak

Uji serentak adalah uji signifikansi model secara keseluruhan atau untuk mengetahui apakah semua variabel prediktor yang dimasukkan kedalam model memberikan pengaruh pada respon secara bersama-sama atau tidak. Perhatikan model regresi 2.3, hipotesis yang digunakan untuk pengujian secara serentak adalah sebagai berikut.

$$\begin{aligned} H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_{p+r} &= 0 \\ H_1 : \text{minimal ada satu } \gamma_j &\neq 0, j = 1, 2, \dots, p+r \end{aligned} \quad (2.10)$$

dengan statistik uji

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \quad (2.11).$$

Daerah penolakan yang digunakan adalah dengan membandingkan nilai  $F_{hitung}$  dengan  $F_{\alpha, (p+r), n-(p+r)-1}$ . Keputusan  $H_0$  ditolak jika  $F_{hitung}$  lebih besar  $F_{tabel}$  atau  $H_0$  ditolak apabila  $p\text{-value} < \alpha$ . Artinya dapat disimpulkan bahwa minimal terdapat satu parameter pada model regresi *spline* yang signifikan (Drapper & Smith, 1992). Perhitungan nilai statistik uji  $F$  didapatkan dari Analisis Ragam (ANOVA) sebagaimana yang ditunjukkan Tabel 2.1.

**Tabel 2.1** Analisis Ragam (ANOVA)

Sumber Variasi	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Rataan Kuadrat (MS)	F <i>hitung</i>
Regresi	$(p + r)$	$\hat{\beta}'\mathbf{X}'y - n\bar{y}^2$	$\frac{SS_{regresi}}{df_{regresi}}$	$\frac{MS_{regresi}}{MS_{residual}}$
Residual	$n - (p + r) - 1$	$y'y - \hat{\beta}'\mathbf{X}'y$	$\frac{SS_{residual}}{df_{residual}}$	
Total	$n - 1$	$y'y - n\bar{y}^2$	-	

### b. Pengujian Secara Individu

Pengujian ini bertujuan untuk mengetahui parameter mana yang berpengaruh dan tidak berpengaruh secara signifikan terhadap model regresi. Hipotesis yang digunakan dalam uji individu adalah sebagai berikut.

$$H_0 : \gamma_j = 0 \quad (2.12)$$

$$H_1 : \gamma_j \neq 0, j = 1, 2, \dots, p+r$$

dengan statistik uji

$$t_{hitung} = \frac{\hat{\gamma}_j}{SE(\hat{\gamma}_j)}, j = 1, 2, \dots, p+r \quad (2.13)$$

dengan  $SE(\hat{y}_j)$  adalah *standart error* dari  $y_i$  dimana nilai ini diperoleh dari  $\sqrt{var(\hat{y}_j)}$  yang dapat diuraikan sebagai berikut.

$$\begin{aligned}
 var(\hat{y}_j) &= var[(X'X)^{-1}X'y] \\
 &= (X'X)^{-1}X' var(y)[(X'X)^{-1}X']' \\
 &= (X'X)^{-1}X' var(y)X(X'X)^{-1} \\
 &= (X'X)^{-1}X'(\sigma^2 I)X(X'X)^{-1} \\
 &= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} \\
 &= \sigma^2(X'X)^{-1}
 \end{aligned} \tag{2.14}$$

Daerah penolakan yang digunakan  $H_0$  ditolak jika  $|t_{hitung}|$  lebih besar  $t_{(\frac{\alpha}{2}, n-(p+r)-1)}$  atau  $H_0$  ditolak apabila  $p\text{-value} < \alpha$ . Artinya dapat disimpulkan bahwa parameter berpengaruh signifikan terhadap model (Drapper & Smith, 1992).

## 2.8 Pengujian Asumsi Residual

Setelah mendapatkan model terbaik dari regresi *spline*, perlu dilakukan pengujian asumsi residual untuk mengetahui apakah residual yang dihasilkan memenuhi asumsi identik, independen dan berdistribusi normal (IIDN). Berikut ini adalah uraian-uraian mengenai asumsi-asumsi yang harus dipenuhi.

### a. Uji Asumsi Identik

Pengujian asumsi identik digunakan untuk mengetahui homogenitas variansi residual. Jika asumsi ini tidak terpenuhi artinya terdapat heteroskedastisitas yang mengakibatkan kerugian bagi efisiensi estimator (Eubank, 1999). Pengujian ini biasanya dilakukan dengan menggunakan uji *glejser*. Uji *glejser* dilakukan dengan meregresikan absolut dari residual dengan variabel prediktornya (Gujarati, 2003). Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{p+r} = 0 \tag{2.15}$$

$$H_1: \text{minimal ada satu } \beta_i \neq 0, i = 1, 2, \dots, p + r$$

dengan statistik uji

$$F_{hitung} = \frac{[\sum_{i=1}^n (|\hat{e}_i| - |\bar{e}|)^2] / ((p+r)-1)}{[\sum_{i=1}^n (|e_i| - |\hat{e}_i|)^2] / (n-(p+r))} \quad (2.16).$$

Daerah penolakan yang digunakan adalah dengan membandingkan nilai  $F_{hitung}$  dengan  $F_{tabel}$ . Keputusan akan bernilai  $H_0$  ditolak jika  $F_{hitung}$  lebih besar dari  $F_{tabel}$  atau  $H_0$  ditolak apabila  $p\text{-value} < \alpha$ . Apabila gagal tolak  $H_0$  maka nilai absolut residual sama dengan nilai  $\beta_0$ , dimana nilai absolut residual merupakan variansi nilai  $y$  dengan  $\hat{y}$ . Jika nilai variansi sama dengan  $\beta_0$  atau konstan, maka variansi di setiap titik  $x$  sama. Artinya dapat disimpulkan bahwa tidak terdapat indikasi adanya kasus heteroskedastisitas pada residual. Solusi untuk mengatasi adanya kasus tersebut adalah melakukan transformasi..

### b. Uji Asumsi Independen

Asumsi independen bertujuan untuk mengetahui terdapat autokorelasi atau tidak. Pengujian autokorelasi dilakukan melalui plot *Autocorrelation Function* (ACF). Berikut ini merupakan formula untuk menghitung ACF (Wei, 2006).

$$\hat{\rho}_l = \frac{\hat{\gamma}_l}{\hat{\gamma}_0} = \frac{\sum_{i=1}^{n-l} (e_i - \bar{e})(e_{i+l} - \bar{e})}{\sum_{i=1}^n (e_i - \bar{e})^2}, \quad l = 1, 2, 3 \dots \quad (2.17)$$

dengan

$\hat{\rho}_l$  = korelasi antara  $e_i$  dan  $e_{i+l}$

$\hat{\gamma}_l$  = kovarian antara  $e_i$  dan  $e_{i+l}$

$\hat{\gamma}_0$  = variansi  $e_i$  = variansi  $e_{i+l}$

$l$  = lag ke-  $l$ .

Interval konfidensi  $(1 - \alpha)$  100% untuk autokorelasi  $\rho_l$  diberikan oleh.

$$-t_{(1-\frac{\alpha}{2}, n-1)} SE(\hat{\rho}_l) \leq \rho_l \leq t_{(\frac{\alpha}{2}, n-1)} SE(\hat{\rho}_l) \quad (2.18)$$

dengan

$$SE(\hat{\rho}_l) = \sqrt{\frac{(1+2 \sum_{m=1}^{l-1} (\hat{\rho}_m))^2}{n}} \quad (2.19).$$

Apabila tidak ada autokorelasi ( $\rho_l$ ) yang keluar dari batas atas maupun batas bawah interval konfidensi maka dapat disimpulkan bahwa asumsi independen telah terpenuhi atau dengan kata lain, tidak ada autokorelasi antar residual. Begitu sebaliknya, bila terdapat autokorelasi ( $\rho_l$ ) yang keluar dari batas interval konfidensi maka dapat disimpulkan asumsi independen tidak terpenuhi. Solusi untuk mengatasi kasus tersebut adalah melakukan differencing.

### c. Uji Asumsi Distribusi Normal

Pengujian asumsi distribusi normal dilakukan untuk mengetahui apakah residual telah berdistribusi normal atau tidak. Untuk mendeteksi apakah residual telah berdistribusi normal dilakukan dengan menggunakan uji *Kolmogoror-Smirnov*. Hipotesis yang digunakan adalah sebagai berikut.

$$H_0 : F(x) = F_0(x) \quad (2.20)$$

$$H_1 : F(x) \neq F_0(x)$$

dengan statistik uji

$$D = maks |F_0(x) - S_N(x)| \quad (2.21)$$

dimana

$F_0(x)$  = fungsi distribusi frekuensi kumulatif

$S_N(x)$  =  $k/N$  adalah fungsi peluang kumulatif yang diobservasi dari suatu sampel random dengan  $N$  observasi dan  $k$  banyak observasi dengan  $k < x$ .

Daerah penolakan yang digunakan adalah  $H_0$  ditolak apabila  $|D| > q_{(1-\alpha)}$  dengan nilai  $q_{(1-\alpha)}$  didapatkan dari tabel Kolmogorov-Smirnov atau  $H_0$  ditolak apabila  $p_{value} < \alpha$ . Solusi untuk mengatasi kasus residual tidak berdistribusi normal ialah menggunakan transformasi *box-cox*.



## 2.9 Indeks Pembangunan Manusia

Menurut UNDP, Indeks Pembangunan Manusia mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup. Sebagai ukuran kualitas hidup, IPM dibangun melalui pendekatan tiga dimensi yaitu, dimensi umur panjang dan sehat, dimensi pengetahuan, dan dimensi kehidupan yang layak (BPS, 2011). Ketiga dimensi tersebut memiliki pengertian sangat luas karena terkait banyak faktor. Untuk mengukur dimensi kesehatan, digunakan angka harapan hidup waktu lahir. Selanjutnya untuk mengukur dimensi pengetahuan digunakan gabungan indikator angka melek huruf dan rata-rata lama sekolah. Adapun untuk mengukur dimensi hidup layak digunakan indikator kemampuan daya beli (*Purchasing Power Parity*). Kemampuan daya beli masyarakat terhadap sejumlah kebutuhan pokok yang dilihat dari rata-rata besarnya pengeluaran perkapita sebagai pendekatan pendapatan yang mewakili capaian pembangunan untuk hidup layak. Formula perhitungan IPM dapat disajikan sebagai berikut.

$$IPM = \frac{1}{3} [X_1 + X_2 + X_3] \quad (2.22)$$

dimana:

$X_1$  = Indeks harapan hidup

$X_2$  = Indeks pendidikan

$X_3$  = Indeks standar hidup layak.

Sebelum menghitung IPM, setiap komponen dari IPM harus terlebih dahulu dihitung indeksnya sehingga bernilai antara 0 (terburuk), dan 1 (terbaik). Untuk memudahkan dalam menganalisa biasanya indeks ini dikalikan 100. Teknik penyusunan indeks tersebut pada dasarnya mengikuti formula sebagai berikut.

$$IPM = \sum_{i=1}^3 li \quad (2.23)$$

dengan

$$li = \frac{X_i - \text{Min } X_i}{\text{Max } X_i - \text{Min } X_i} \quad (2.24)$$

dimana:

$li$  = Indeks komponen IPM ke- $i$ , dimana  $i = 1,2,3$

$X_i$  = Nilai indikator komponen ke- $i$

$\text{Max } X_i$  = Nilai maksimum  $X_i$

$\text{Min } X_i$  = Nilai minimum  $X_i$ .

Untuk menghitung indeks masing-masing komponen IPM digunakan batas maksimum dan minimum pada Tabel 2.1 sebagai berikut

**Tabel 2.2** Nilai Maksimum dan Minimum dari setiap Komponen IPM

Komponen IPM	Maksimum	Minimum	Keterangan
Angka Harapan Hidup (tahun)	85	25	Standar UNDP
Angka Melek Huruf (persen)	100	0	Standar UNDP
Rata-rata Lama Sekolah (tahun)	15	0	
Daya Beli (ribu rupiah PPP)	737.730 <sup>a</sup>	300.000(1996) 360.000 <sup>b</sup> (1999, dst)	Pengeluaran per Kapita Riil disesuaikan

Indeks Pembangunan Manusia (IPM) merupakan indeks komposit yang dihitung sebagai rata-rata sederhana dari tiga indeks dasar yaitu indeks harapan hidup, indeks pendidikan dan indeks standar hidup layak. Berikut ini adalah komponen pembentukan Indeks Pembangunan Manusia

a. Angka Harapan Hidup

Angka Harapan Hidup adalah rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang selama hidup atau perkiraan lama hidup rata-rata penduduk dengan asumsi tidak ada perubahan pola mortalitas menurut umur. Ada dua jenis data yang digunakan dalam perhitungan Angka Harapan Hidup yaitu Angka

Lahir Hidup (ALH) dan Anak Masih Hidup (AMH). (Preston dkk, 2004).

b. Tingkat Pendidikan

Untuk mengukur dimensi pengetahuan penduduk digunakan dua indikator, yaitu rata-rata lama sekolah dan angka melek huruf. Rata-rata lama sekolah menggambarkan jumlah tahun yang digunakan oleh penduduk usia 15 tahun keatas dalam menjalani pendidikan formal. Sedangkan angka melek huruf adalah persentase usia 15 tahun keatas yang dapat membaca dan menulis huruf latin dan tau huruf lainnya.

c. Standar Hidup Layak

Dimensi dari ukuran kualitas hidup manusia adalah standar hidup layak. Dalam cakupan lebih luas standar hidup layak menggambarkan tingkat kesejahteraan yang dinikmati oleh penduduk sebagai dampak semakin membaiknya ekonomi. BPS mengukur standar hidup layak menggunakan rata-rata pengeluaran perkapita riil.

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Sumber Data**

Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari publikasi Statistika Daerah Jawa Tengah 2015, Jawa Tengah dalam angka 2015 dan website Badan Pusat Statistika Jawa Tengah. Data yang digunakan ialah data pada tahun 2013 dengan banyaknya observasi yang digunakan adalah 35 kabupaten/kota di Jawa Tengah.

#### **3.2 Variabel Penelitian**

Variabel respon yang digunakan pada penelitian ini adalah Indeks Pembangunan Manusia (IPM) di 35 kabupaten/kota di provinsi Jawa Tengah. Variabel – variabel yang diduga berpengaruh didapatkan dari penelitian sebelumnya yang diuraikan dalam Tabel 3.1.

**Tabel 3.1** Variabel Penelitian

<b>Variabel</b>	<b>Keterangan</b>
$y$	Indeks Pembangunan Manusia
$x_1$	Tingkat Partisipasi Angkatan Kerja
$x_2$	Rasio Sekolah-Siswa
$x_3$	Kepadatan Penduduk
$x_4$	Angka Kesakitan
$x_5$	PDRB/1juta

Berikut ini merupakan keterangan mengenai variabel penelitian.

- a. Variabel  $y$  merupakan variabel respon yang menyatakan Indeks Pembangunan Manusia. Nilai IPM ini didapatkan dari:

$$IPM = \frac{1}{3} [X_1 + X_2 + X_3]$$

dimana:

$X_1$  = Indeks harapan hidup

$X_2$  = Indeks pendidikan

$X_3$  = Indeks standar hidup layak

- b. Variabel  $x_1$  merupakan variabel yang menyatakan Tingkat Partisipasi Angkatan Kerja (TPAK). TPAK merupakan persentase jumlah angkatan kerja terhadap penduduk usia kerja. Nilai TPAK diperoleh dari:

$$TPAK = \frac{\text{jumlah angkatan kerja}}{\text{jumlah penduduk usia 15 tahun keatas}} \times 100$$

- c. Variabel  $x_2$  merupakan variabel yang menyatakan Rasio Sekolah-Siswa pada SMA. Perbandingan jumlah siswa dengan jumlah sekolah pada jenjang pendidikan SMA untuk mengetahui rata-rata jumlah siswa yang dapat belajar di suatu sekolah. Nilai Rasio Sekolah-Siswa didapatkan dari:

$$\text{Rasio Sekolah} - \text{Siswa} = \frac{\text{jumlah Siswa SMA}}{\text{jumlah sekolah SMA}}$$

- d. Variabel  $x_3$  merupakan variabel yang menyatakan Kepadatan Penduduk. Kepadatan Penduduk merupakan banyaknya penduduk suatu satuan luas. Nilai Kepadatan Penduduk diperoleh dari:

$$KP = \frac{\text{Jumlah Penduduk}}{\text{Luas Wilayah (km}^2\text{)}}$$

- e. Variabel  $x_4$  merupakan variabel yang menyatakan Angka Kesakitan. Angka Kesakitan merupakan persentase penduduk yang mengalami gangguan kesehatan dan terganggu aktifitasnya sehari-hari yang terjadi selama satu bulan sebelum pencacahan.

- f. Variabel  $x_5$  merupakan variabel PDRB/1juta. PDRB merupakan total nilai produksi barang dan jasa yang diproduksi di wilayah (regional) tertentu dalam waktu tertentu (satu tahun).

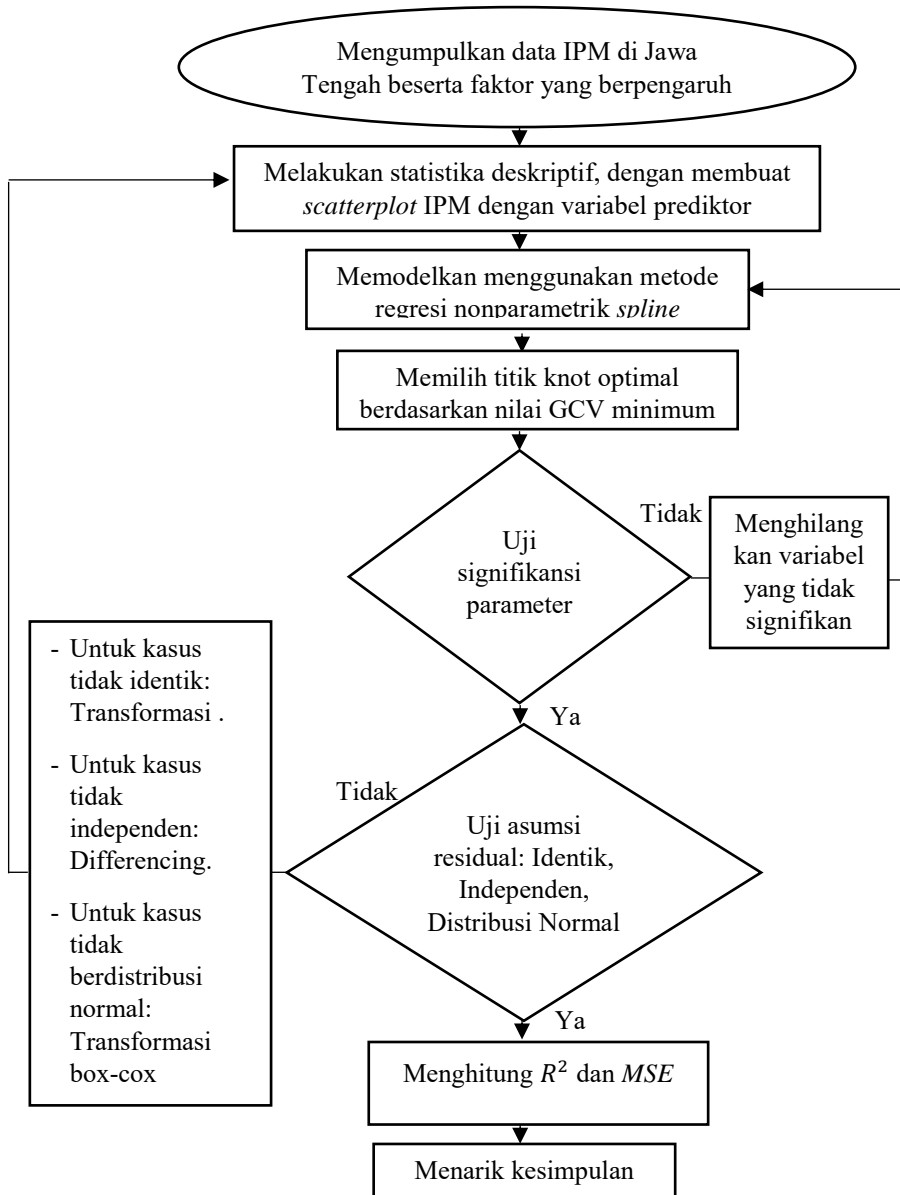
### 3.3 Langkah Penelitian

Langkah-langkah dalam penelitian ini adalah sebagai berikut.

1. Mengumpulkan data Indeks Pembangunan Manusia di Jawa Tengah beserta variabel-variabel yang berpengaruh.
2. Mendeskripsikan karakteristik dari data Indeks Pembangunan Manusia di Jawa Tengah beserta variabel-variabel yang berpengaruh, menggunakan *scatterplot* antara Indeks Pembangunan Manusia di Jawa Tengah dengan variabel-variabel prediktornya untuk mengetahui bentuk pola data. Apabila pola data tidak mengikuti pola tertentu maka metode yang digunakan adalah regresi nonparametrik *spline*.
3. Memodelkan data dengan pendekatan *spline* dengan satu, dua, tiga titik knot dan kombinasi knot.
4. Memilih titik knot optimal berdasarkan nilai *Generalized Cross Validation* (GCV) yang paling minimum.
5. Mendapatkan model regresi *spline* dengan titik knot optimal.
6. Menguji signifikansi parameter regresi *spline* secara serentak dan parsial.
7. Mendeteksi asumsi residual model *spline* terbaik. Apabila residual model *spline* tidak memenuhi asumsi, maka dilakukan transformasi data. Kemudian, memulai kembali dari langkah (2).
8. Menghitung nilai koefisien determinasi  $R^2$
9. Menarik kesimpulan.

### 3.4 Diagram Alir

Dari langkah analisis disajikan secara ringkas dalam diagram alir yang dijelaskan pada Gambar 3.1.



**Gambar 3.1** Diagram Alir

## **BAB IV**

### **ANALISIS DAN PEMBAHASAN**

#### **4.1 Karakteristik Indeks Pembangunan Manusia di Jawa Tengah**

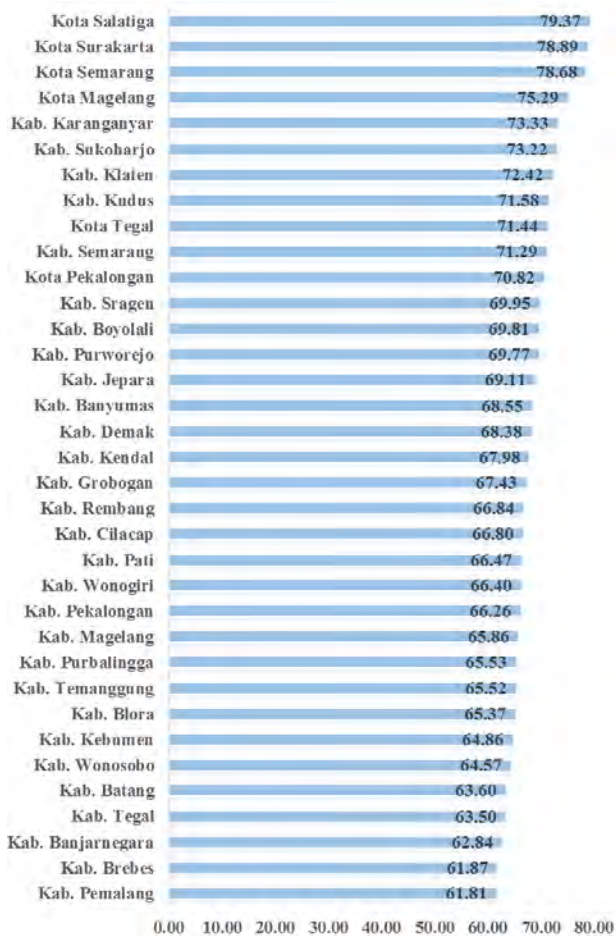
Menurut UNDP, IPM mengukur pencapaian hasil pembangunan dari suatu daerah/wilayah dalam tiga dimensi dasar pembangunan yaitu kesehatan, pendidikan, dan standar hidup layak. Setiap tahunnya nilai IPM Jawa Tengah mengalami peningkatan, pada tahun 2013 Jawa Tengah memiliki nilai IPM sebesar 68,02. Gambar 4.1 menyajikan diagram batang nilai IPM di setiap kabupaten/kota provinsi Jawa Tengah sesuai dengan data di Lampiran 1.

Berdasarkan diagram batang yang telah diurutkan dari nilai IPM terendah hingga tertinggi pada Gambar 4.1 diketahui bahwa di provinsi Jawa Tengah nilai IPM terendah yaitu Kabupaten Pemalang dengan nilai IPM sebesar 61,81. Dilihat dari angka IPM tersebut, berdasarkan kategori IPM yang dikeluarkan oleh PBB, Kabupaten Pemalang termasuk ke dalam IPM menengah bawah (50,0 sampai 65,9). Nilai IPM tertinggi di Jawa Tengah berada pada Kota Salatiga dengan nilai sebesar 79,37. Kota Salatiga termasuk ke dalam IPM menengah atas (66,0 sampai 79,9).

Bila dilihat dari kategori IPM terdapat 11 kabupaten yang masih berada di kategori menengah bawah yaitu Kab. Pemalang, Kab. Brebes, Kab. Banjarnegara, Kab. Tegal, Kab. Batang, Kab. Wonosobo, Kab. Kebumen, Kab. Blora, Kab. Temanggung, Kab. Purbalingga, Kab. Magelang. Nilai IPM Kab. Magelang sebesar 65,86. Sisanya ialah kategori IPM menengah atas dengan 18 kabupaten dan 5 kota. Walaupun sebagian besar kabupaten/kota provinsi Jawa Tengah sudah berada dalam kategori menengah atas, namun belum ada yang mampu menembus kategori tinggi. Dikarenakan pemerintah Indonesia menargetkan nilai IPM berada pada kategori tinggi di setiap provinsi yang ada di Indonesia, maka



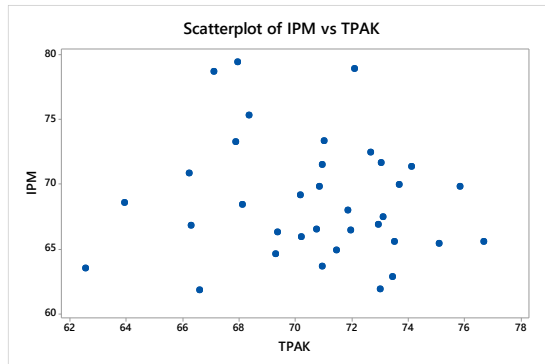
IPM di Jawa Tengah perlu ditingkatkan. Sehingga IPM di Jawa Tengah bisa mencapai kategori tinggi dengan memperhatikan faktor-faktor yang memengaruhi nilai IPM.



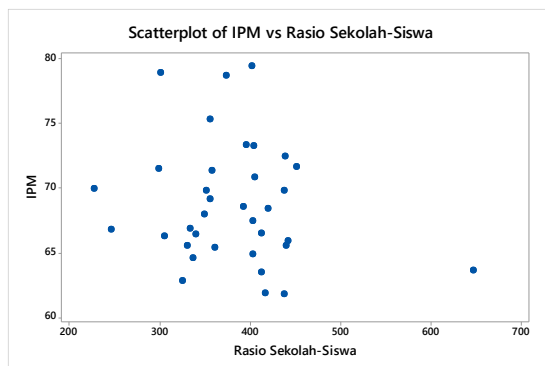
**Gambar 4.1** Indeks Pembangunan Manusia di Kabupaten/Kota Provinsi Jawa Tengah

## 4.2 Analisis Faktor-faktor yang Diduga Memengaruhi Indeks Pembangunan Manusia

Selanjutnya, dilakukan analisis deskriptif dari variabel-variabel yang diduga berpengaruh terhadap IPM di provinsi Jawa Tengah. Salah satunya ialah membuat *scatterplot* antara variabel respon dengan variabel prediktor agar mengetahui pola hubungan IPM dengan masing-masing variabel prediktor. Pada Gambar 4.2 diketahui variabel IPM dengan tingkat partisipasi angkatan kerja tidak membentuk suatu pola tertentu, sehingga estimasi model menggunakan pendekatan regresi nonparametrik.

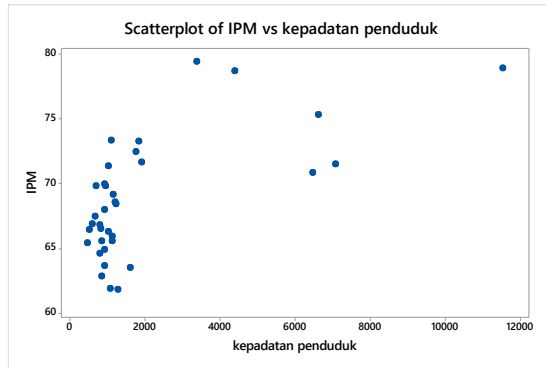


**Gambar 4.2** Pola Antara IPM dengan TPAK



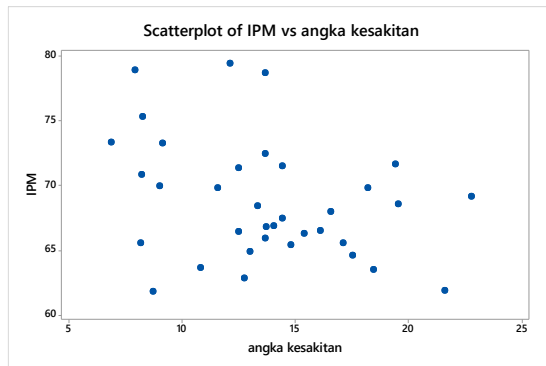
**Gambar 4.3** Pola Antara IPM dengan Rasio Sekolah-Siswa

Dapat dilihat bahwa Gambar 4.3 merupakan pola hubungan IPM dengan rasio sekolah-siswa yang tidak membentuk suatu pola tertentu, sehingga akan dilakukan estimasi model menggunakan regresi nonparametrik



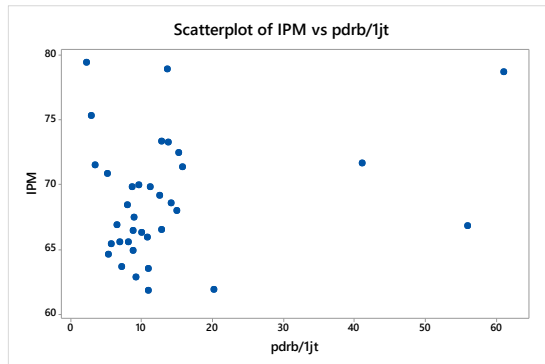
**Gambar 4.4** Pola Antara IPM dengan Kepadatan Penduduk

Berdasarkan Gambar 4.4 menunjukkan bahwa pola hubungan IPM dengan kepadatan penduduk tidak membentuk suatu pola tertentu, sehingga estimasi model menggunakan pendekatan regresi nonparametrik.



**Gambar 4.5** Pola Antara IPM dengan Angka Kesakitan

Berdasarkan Gambar 4.5 diketahui bahwa antara IPM dengan jumlah dan angka kesakitan memberikan pola tidak membentuk suatu pola tertentu, maka estimasi model menggunakan regresi nonparametrik.



**Gambar 4.6** Pola Antara IPM dengan PDRB/1juta

Dapat dilihat bahwa Gambar 4.6 merupakan pola hubungan IPM dengan PDRB/1juta yang tidak membentuk suatu pola tertentu, sehingga akan dilakukan estimasi model menggunakan regresi nonparametrik. Pada penelitian ini metode yang baik digunakan ialah pendekatan regresi nonparametrik *spline*. Berikut ini adalah karakteristik dari kelima variabel yang diduga memengaruhi IPM provinsi Jawa Tengah.

**Tabel 4.1** Statistika Deskriptif Variabel Penelitian

Variabel	Mean	St. Deviasi	Minimum	Maximum
$X_1$	70,672	3,264	62,56	76,7
$X_2$	380,4	72,1	227,1	646,9
$X_3$	1985	2403	471	11534
$X_4$	13,719	4,053	6,880	22,790
$X_5$	13,55	13,05	2,28	61,09

Variabel  $X_1$  merupakan variabel tingkat partisipasi angkatan kerja yang diduga memengaruhi IPM. Semakin tinggi tingkat partisipasi angkatan kerja menunjukkan bahwa semakin tinggi pula pasokan tenaga kerja yang tersedia untuk memproduksi barang dan jasa dalam suatu perekonomian. Pada Tabel 4.1 diketahui bahwa variabel tingkat partisipasi angkatan kerja memiliki rata-rata sebesar 70,672 persen dengan standar deviasi 3,264. Nilai tertinggi terdapat di Kabupaten Temanggung yaitu sebesar 76,7 persen sedangkan tingkat partisipasi angkatan kerja terendah terdapat di Kabupaten Tegal dengan nilai sebesar 62,56 persen.

Variabel  $X_2$  merupakan variabel rasio sekolah-siswa yang diduga memengaruhi IPM. Variabel ini menggambarkan dari dimensi pengetahuan yang terdapat didimensi IPM. Variabel ini mengetahui rata-rata jumlah siswa yang dapat belajar di suatu sekolah. Pada Tabel 4.1 diketahui bahwa variabel rasio sekolah-siswa memiliki nilai rata-rata sebesar 380,4 dengan standar deviasi sebesar 72,1. Nilai tertinggi bernilai 646,9 terdapat di Kabupaten Batang sedangkan nilai terendah yaitu 227,1 terdapat di Kabupaten Sragen.

Variabel  $X_3$  merupakan variabel kepadatan penduduk yang diduga memengaruhi IPM. Angka kepadatan penduduk menunjukkan rata-rata jumlah penduduk tiap 1 kilometer persegi. Semakin besar angka kepadatan penduduk menunjukan bahwa semakin padat penduduk yang mendiami wilayah tersebut. Dalam Tabel 4.1 diketahui bahwa variabel kepadatan penduduk memiliki rata-rata sebesar 1985 dengan standar deviasi 2403. Nilai tertinggi terdapat di Kota Surakarta yaitu sebesar 11534 sedangkan kepadatan penduduk terendah terdapat di Kabupaten Blora dengan nilai sebesar 471.

Variabel  $X_4$  merupakan variabel angka kesakitan yang diduga memengaruhi IPM. Variabel ini menggambarkan dari dimensi kesehatan yang terdapat didimensi IPM. Variabel ini dapat mengukur tingkat kesehatan masyarakat secara umum yang dilihat dari adanya keluhan yang mengindikasikan terkena suatu penyakit

tertentu. Pada Tabel 4.1 diketahui bahwa variabel angka kesakitan memiliki nilai rata-rata sebesar 13,719 dengan standar deviasi sebesar 4,053. Nilai tertinggi bernilai 22,790 terdapat di Kabupaten Jepara sedangkan nilai terendah yaitu 6,880 terdapat di Kabupaten Karanganyar.

Variabel  $X_5$  merupakan variabel PDRB/1juta yang diduga memengaruhi IPM. Dimensi yang terdapat dalam IPM yaitu standar hidup layak menggambarkan tingkat kesejahteraan yang dinikmati oleh penduduk sebagai dampak semakin membaiknya ekonomi, hal tersebut dapat diukur dari PDRB/1juta. Dalam Tabel 4.1 variabel PDRB/1juta memiliki rata-rata sebesar 13,55 persen dengan standar deviasi sebesar 13,05. Nilai tertinggi terdapat di Kota Semarang yaitu sebesar 61,9 sedangkan PDRB/1juta terendah terdapat di Kota Salatiga dengan nilai sebesar 2,28.

### **4.3 Regresi Nonparametrik**

Model regresi nonparametrik *spline* yang digunakan sebagai pendekatan dalam menyatakan hubungan antara IPM dengan lima variabel yang diduga memengaruhinya adalah regresi nonparametrik *spline* dengan 1 titik knot, 2 titik knot, 3 titik knot, dan kombinasi titik knot.

#### **4.3.1 Pemilihan Titik Knot Optimum**

Titik knot merupakan titik perpaduan dimana terjadi perubahan pola data. Model regresi nonparametrik *spline* terbaik didapatkan dari titik knot yang optimal. Untuk menentukan titik knot yang optimal, digunakan metode GCV (*Generalized Cross Validation*). Titik knot optimal diperoleh dari nilai GCV yang minimum, nilai GCV didapatkan dari perhitungan dengan menggunakan rumus pada persamaan 2.5. Berikut akan ditampilkan nilai GCV dengan 1 titik knot, 2 titik knot, 3 titik knot, dan kombinasi titik knot.

## 1. Pemilihan Titik Knot dengan Satu Titik Knot

Estimasi model regresi nonparametrik *spline* dengan satu titik knot pada nilai IPM Provinsi Jawa Tengah adalah sebagai berikut.

$$\hat{y} = \hat{\gamma}_0 + \hat{\gamma}_1 x_1 + \hat{\gamma}_2 (x_1 - K_1)_+^1 + \hat{\gamma}_3 x_2 + \hat{\gamma}_4 (x_2 - K_2)_+^1 + \hat{\gamma}_5 x_3 + \hat{\gamma}_6 (x_3 - K_3)_+^1 + \hat{\gamma}_7 x_4 + \hat{\gamma}_8 (x_4 - K_4)_+^1 + \hat{\gamma}_9 x_5 + \hat{\gamma}_{10} (x_5 - K_5)_+^1$$

Tabel 4.2 merupakan sepuluh nilai GCV disekitar nilai GCV yang paling minimum untuk model regresi nonparametrik *spline*.

**Tabel 4.2** Nilai GCV dengan Satu Titik Knot

GCV	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
15,20	64,29	278,47	1825,65	8,83	9,48
14,48	63,14	244,20	922,55	7,53	4,68
15,34	63,43	252,77	1148,33	7,85	5,88
15,07	63,71	261,34	1374,10	8,18	7,08
15,34	64,00	269,91	1599,88	8,50	8,28
<b>12,51</b>	<b>62,85</b>	<b>235,64</b>	<b>696,78</b>	<b>7,20</b>	<b>3,48</b>
15,06	64,58	287,04	2051,43	9,15	10,68
14,48	64,87	295,61	2277,20	9,48	11,88
13,67	65,16	304,18	2502,98	9,80	13,08
13,07	65,45	312,74	2728,76	10,13	14,28

Pada Tabel 4.2 diketahui bahwa nilai GCV pada 1 titik knot yang paling minimum adalah 12,51 dengan titik knot optimum yaitu variabel tingkat partisipasi angkatan kerja (X<sub>1</sub>) pada titik 62,85; variabel rasio sekolah-siswa (X<sub>2</sub>) pada titik 235,64; variabel kepadatan penduduk (X<sub>3</sub>) pada titik 696,78; variabel angka kesakitan (X<sub>4</sub>) pada titik 7,20 serta variabel PDRB/1juta (X<sub>5</sub>) pada titik 3,48

## 2. Pemilihan Titik Knot dengan Dua Titik Knot

Setelah dilakukan pemilihan titik knot dengan satu titik knot, selanjutnya akan dilakukan pemilihan titik knot optimal menggunakan dua titik knot. Estimasi model regresi nonparametrik

*spline* dengan dua titik knot pada nilai IPM Provinsi Jawa Tengah adalah sebagai berikut.

$$\begin{aligned}\hat{y} = & \hat{\gamma}_0 + \hat{\gamma}_1 x_1 + \hat{\gamma}_2 (x_1 - K_1)_+^1 + \hat{\gamma}_3 (x_1 - K_2)_+^1 + \hat{\gamma}_4 x_2 + \\ & \hat{\gamma}_5 (x_2 - K_3)_+^1 + \hat{\gamma}_6 (x_2 - K_4)_+^1 + \hat{\gamma}_7 x_3 + \hat{\gamma}_8 (x_3 - K_5)_+^1 + \\ & \hat{\gamma}_9 (x_3 - K_6)_+^1 + \hat{\gamma}_{10} x_4 + \hat{\gamma}_{11} (x_4 - K_7)_+^1 + \hat{\gamma}_{12} (x_4 - K_8)_+^1 + \\ & \hat{\gamma}_{13} x_5 + \hat{\gamma}_{14} (x_5 - K_9)_+^1 + \hat{\gamma}_{15} (x_5 - K_{10})_+^1\end{aligned}$$

Tabel 4.3 merupakan sepuluh nilai GCV minimum untuk model regresi nonparametrik *spline* menggunakan dua titik knot.

**Tabel 4.3** Nilai GCV dengan Dua Titik Knot

GCV	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
15,59	66,31	338,45	3406,08	11,10	17,89
	75,55	612,60	10630,90	21,49	56,29
14,59	66,31	338,45	3406,08	11,10	17,89
	75,83	621,17	10856,67	21,82	57,49
14,53	66,31	338,45	3406,08	11,10	17,89
	76,12	629,73	11082,45	22,14	58,69
14,56	66,31	338,45	3406,08	11,10	17,89
	76,41	638,30	11308,22	22,47	59,89
12,84	66,31	338,45	3406,08	11,10	17,89
	76,70	646,87	11534,00	22,79	61,09
<b>9,68</b>	<b>66,60</b>	<b>347,01</b>	<b>3631,86</b>	<b>11,43</b>	<b>19,09</b>
	<b>66,89</b>	<b>355,58</b>	<b>3857,63</b>	<b>11,75</b>	<b>20,29</b>
10,34	66,60	347,01	3631,86	11,43	19,09
	67,18	364,15	4083,41	12,08	21,49
11,48	66,60	347,01	3631,86	11,43	19,09
	67,47	372,71	4309,18	12,40	22,69
12,21	66,60	347,01	3631,86	11,43	19,09
	67,75	381,28	4534,96	12,72	23,89
13,23	66,60	347,01	3631,86	11,43	19,09
	68,04	389,85	4760,73	13,05	25,09



Berdasarkan Tabel 4.3 diketahui bahwa nilai GCV minimum yang diperoleh dari pemodelan regresi nonparametrik *spline* dengan dua titik knot adalah 9,68. Dengan titik knot optimum yaitu variabel  $X_1$  pada titik 66,60 dan 66,89; variabel  $X_2$  pada titik 347,01 dan 355,58; variabel  $X_3$  pada titik 3631,86 dan 3857,63; variabel  $X_4$  pada titik 11,43 dan 11,75; variabel  $X_5$  pada titik 19,09 dan 20,29. Nilai GCV dengan 2 titik knot bernilai lebih kecil dari 1 titik knot sehingga mode dengan 2 titik knot lebih baik daripada dengan 1 titik knot. Akan tetapi masih harus dicobakan dengan menggunakan 3 titik knot dan kombinasi titik knot untuk mendapatkan kemungkinan model lebih baik.

### 3. Pemilihan Titik Knot dengan Tiga Titik Knot

Estimasi model regresi nonparametrik *spline* dengan tiga titik knot pada nilai IPM Provinsi Jawa Tengah adalah sebagai berikut.

$$\begin{aligned}\hat{y} = & \hat{\gamma}_0 + \hat{\gamma}_1 x_1 + \hat{\gamma}_2 (x_1 - K_1)_+^1 + \hat{\gamma}_3 (x_1 - K_2)_+^1 + \hat{\gamma}_4 (x_1 - K_3)_+^1 + \\ & \gamma_5 x_2 + \hat{\gamma}_6 (x_2 - K_4)_+^1 + \hat{\gamma}_7 (x_2 - K_5)_+^1 + \hat{\gamma}_8 (x_2 - K_6)_+^1 + \hat{\gamma}_9 x_3 + \\ & \hat{\gamma}_{10} (x_3 - K_7)_+^1 + \hat{\gamma}_{11} (x_3 - K_8)_+^1 + \hat{\gamma}_{12} (x_3 - K_9)_+^1 + \hat{\gamma}_{13} x_4 + \\ & \hat{\gamma}_{14} (x_4 - K_{10})_+^1 + \hat{\gamma}_{15} (x_4 - K_{11})_+^1 + \hat{\gamma}_{16} (x_4 - K_{12})_+^1 + \hat{\gamma}_{17} x_5 + \\ & \hat{\gamma}_{18} (x_5 - K_{13})_+^1 + \hat{\gamma}_{19} (x_5 - K_{14})_+^1 + \hat{\gamma}_{20} (x_5 - K_{15})_+^1\end{aligned}$$

Berikut ini merupakan lima nilai diantara nilai GCV minimum yang diperoleh dari regresi nonparametrik *spline* dengan tiga titik knot yang disajikan dalam Tabel 4.4.

Berdasarkan Tabel 4.4 diketahui bahwa nilai GCV minimum yang diperoleh dari pemodelan regresi nonparametrik *spline* dengan tiga titik knot adalah 6,36. Dengan titik knot optimum yaitu variabel  $X_1$  pada titik 65,73, 66,60, dan 66,89; variabel  $X_2$  pada titik 321,31, 347,01, dan 355,58; variabel  $X_3$  pada titik 2954,53, 3631,86, dan 3857,63; variabel  $X_4$  pada titik 10,45, 11,43, dan 11,75; variabel  $X_5$  pada titik 15,48, 19,09, dan 20,29. Nilai GCV dengan 3 titik knot bernilai lebih kecil dari 1 titik knot dan 2 titik knot sebelumnya sehingga model dengan 3 titik knot lebih baik. Akan tetapi masih harus dicobakan lagi menggunakan kombinasi titik knot untuk mendapatkan kemungkinan model yang lebih baik

**Tabel 4.4** Nilai GCV dengan Tiga Titik Knot

GCV	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$
11,86	65,73	321,31	2954,53	10,45	15,48
	66,31	338,45	3406,08	11,10	17,89
	76,41	638,30	11308,22	22,47	59,89
<b>6,36</b>	<b>65,73</b>	<b>321,31</b>	<b>2954,53</b>	<b>10,45</b>	<b>15,48</b>
	<b>66,60</b>	<b>347,01</b>	<b>3631,86</b>	<b>11,43</b>	<b>19,09</b>
	<b>66,89</b>	<b>355,58</b>	<b>3857,63</b>	<b>11,75</b>	<b>20,29</b>
7,33	65,73	321,31	2954,53	10,45	15,48
	66,60	347,01	3631,86	11,43	19,09
	67,18	364,15	4083,41	12,08	21,49
7,27	65,73	321,31	2954,53	10,45	15,48
	66,60	347,01	3631,86	11,43	19,09
	67,47	372,71	4309,18	12,40	22,69
6,92	65,73	321,31	2954,53	10,45	15,48
	66,60	347,01	3631,86	11,43	19,09
	67,75	381,28	4534,96	12,72	23,89

#### 4. Pemilihan Titik Knot dengan Kombinasi Titik Knot

Pemilihan titik knot optimal dengan 1 titik knot, 2 titik knot serta 3 titik knot telah dilakukan. Selanjutnya, akan dilakukan pemilihan titik knot optimal dengan kombinasi titik knot karena terdapat kemungkinan bahwa setiap pola data memiliki jumlah titik knot optimal yang berbeda-beda. Berikut ini merupakan lima nilai diantara nilai GCV minimum yang disajikan dalam Tabel 4.5.

Berdasarkan Tabel Tabel 4.7 diketahui bahwa nilai GCV minimum dari kombinasi knot adalah 4,80. Nilai GCV tersebut dihasilkan apabila digunakan kombinasi knot (3,3,2,1,2). Titik knot yang digunakan adalah 65,73, 66,60, dan 66,89 untuk  $X_1$ ; 321,31, 347,01, dan 355,58 untuk  $X_2$ ; 3631,86 dan 3857,63 untuk  $X_3$ ; 7,20 untuk  $X_4$ ; dan 19,09 dan 20,29 untuk  $X_5$ . Selanjutnya akan

dilakukan pemilihan model terbaik berdasarkan nilai GCV yang paling minimum.

**Tabel 4.5** Nilai GCV dengan Kombinasi Knot

GCV	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
7,49	65,73	321,31	3631,86	7,20	3,48
	66,60	347,01	3857,63		
	66,89	355,58			
<b>4,80</b>	<b>65,73</b>	<b>321,31</b>	<b>3631,86</b>	<b>7,20</b>	<b>19,09</b>
	<b>66,60</b>	<b>347,01</b>	<b>3857,63</b>		<b>20,29</b>
	<b>66,89</b>	<b>355,58</b>			
4,91	65,73	321,31	3631,86	7,20	15,48
	66,60	347,01	3857,63		19,09
	66,89	355,58			20,29
9,31	65,73	321,31	3631,86	11,43	3,48
	66,60	347,01	3857,63	11,75	
	66,89	355,58			
6,91	65,73	321,31	3631,86	11,43	19,09
	66,60	347,01	3857,63	11,75	20,29
	66,89	355,58			

#### 4.3.2 Pemilihan Model Terbaik

Setelah didapatkan nilai GCV minimum pada pemilihan titik knot optimal dengan 1 titik knot, 2 titik knot, 3 titik knot dan kombinasi knot, selanjutnya dilakukan pemilihan model terbaik. Pemilihan model terbaik dilakukan dengan membandingkan nilai dari GCV yang paling minimum. Tabel 4.6 merupakan tabel yang menunjukkan nilai GCV terkecil dari masing-masing model.

Berdasarkan Tabel 4.6 diketahui bahwa pemodelan yang menghasilkan nilai GCV paling minimum merupakan pemodelan regresi nonparametrik *spline* dengan menggunakan kombinasi knot (3,3,2,1,2). Oleh karena itu, diputuskan bahwa model terbaik yang

akan dipilih adalah model regresi nonparametrik *spline* dengan menggunakan kombinasi knot 3,3,2,1,2.

**Tabel 4.6** Perbandingan Nilai GCV Minimum

Model	GCV Minimum
1 Titik Knot	12,51
2 Titik Knot	9,68
3 Titik Knot	6,36
Kombinasi Knot (3,3,2,1,2)	4,80

Pemodelan regresi nonparametrik *spline* menggunakan titik knot optimal diperoleh dari nilai GCV paling minimum. Model regresi nonparametrik *spline* terbaik dituliskan sebagai berikut.

$$\begin{aligned}\hat{y} = & 3,699 + 2,561 X_1 - 14,949 (X_1 - 65,73)_+^1 + \\ & 39,517 (X_1 - 66,60)_+^1 - 27,005 (X_1 - 66,89)_+^1 - \\ & 0,057 X_2 + 0,279(X_2 - 321,31)_+^1 - \\ & 0,347(X_2 - 347,01)_+^1 + 0,107(X_2 - 355,58)_+^1 + \\ & 0,005 X_3 - 0,047(X_3 - 3631,86)_+^1 + \\ & 0,043 (X_3 - 3857,63)_+^1 - 12,918 X_4 + \\ & 12,812 (X_4 - 7,20)_+^1 + 0,153 X_5 - \\ & 6,385 (X_5 - 19,09)_+^1 + 6,524(X_5 - 20,29)_+^1\end{aligned}$$

#### 4.3.3 Pengujian Signifikansi Parameter Model Regresi Nonparametrik *Spline*

Pengujian parameter dilakukan untuk mengetahui apakah parameter yang telah didapatkan dari hasil pemodelan dengan regresi nonparametrik *spline* berpengaruh secara signifikan terhadap variabel IPM atau tidak. Pengujian dilakukan dengan dua tahap yaitu pengujian parameter secara serentak dan individu. Apabila pada pengujian secara serentak, parameter menunjukkan pengaruh yang signifikan terhadap IPM maka dilanjutkan pada pengujian secara individu untuk mengetahui pengaruh signifikansi dari tiap-tiap parameter terhadap IPM.

## 1. Uji Serentak

Uji serentak dilakukan pada parameter model regresi terhadap variabel IPM secara bersama-sama atau serentak. Hipotesis Yng digunakan untuk pengujian secara serentak terdapat pada persamaan 2.10. Hasil pengujian secara serentak ditampilkan pada Tabel 4.7.

**Tabel 4.7** Hasil Uji Serentak

Sumber	df	SS	MS	Fhit	<i>p-value</i>
Regresi	16	672,7788	42,04868	15,28829	2,448358e-07
Error	18	49,50693	2,750385		
Total	34	722,2857			

Berdasarkan Tabel 4.7 diketahui bahwa nilai  $p\text{-value}$  sebesar 2.448358e-07. Dengan taraf signifikansi ( $\alpha$ ) sebesar 5% maka didapatkan keputusan tolak  $H_0$  karena nilai  $p\text{-value} < \alpha$  (0,05) dan nilai  $F_{hitung} > F_{(0,05,16,18)}$  yaitu  $15,28829 > 2,249587$ , sehingga disimpulkan bahwa minimal terdapat satu parameter pada model yang signifikan.

## 2. Uji Individu

Uji individu dilakukan untuk mengetahui parameter mana saja yang berpengaruh signifikan terhadap model regresi. Hipotesis yang digunakan dalam uji individu terdapat pada persamaan 2.12. Hasil dari pengujian parameter model regresi secara individu ditampilkan pada Tabel 4.8.

Berdasarkan Tabel 4.8 diketahui bahwa dengan taraf signifikansi ( $\alpha$ ) sebesar 5% dari 17 parameter menunjukkan bahwa semua parameter signifikan terhadap model. Parameter yang signifikan ialah variabel tingkat partisipasi angkatan kerja ( $X_1$ ), rasio sekolah-siswa ( $X_2$ ), kepadatan penduduk ( $X_3$ ), angka kesakitan ( $X_4$ ), dan PDRB/1juta ( $X_5$ ).

**Tabel 4.8** Hasil Uji Individu

Var	Parameter	Estimasi	$p\text{-value}$	Keputusan	Kesimpulan
Const	$\gamma_0$	3,699	0,001	Tolak	Signifikan
$X_1$	$\gamma_1$	2,561	0,000	Tolak	Signifikan
	$\gamma_2$	-14,949	0,000	Tolak	
	$\gamma_3$	39,517	0,000	Tolak	
	$\gamma_4$	-27,005	0,000	Tolak	
$X_2$	$\gamma_5$	-0,057	0,018	Tolak	Signifikan
	$\gamma_6$	0,279	0,012	Tolak	
	$\gamma_7$	-0,347	0,211	Gagal	
	$\gamma_8$	0,107	0,594	Gagal	
$X_3$	$\gamma_9$	0,005	0,000	Tolak	Signifikan
	$\gamma_{10}$	-0,047	0,000	Tolak	
	$\gamma_{11}$	0,043	0,001	Tolak	
$X_4$	$\gamma_{12}$	-12,918	0,023	Tolak	Signifikan
	$\gamma_{13}$	12,812	0,025	Tolak	
$X_5$	$\gamma_{14}$	0,153	0,149	Gagal	Signifikan
	$\gamma_{15}$	-6,385	0,003	Tolak	
	$\gamma_{16}$	6,524	0,002	Tolak	

#### 4.3.4 Pengujian Asumsi Residual

Pengujian asumsi residual dilakukan untuk mengetahui kelayakan suatu model regresi apakah residual yang telah terbentuk memenuhi asumsi identik, independen, dan berdistribusi normal (IIDN) atau tidak. Apabila suatu model regresi dengan kriteria model terbaik dan parameter signifikan namun tidak memenuhi asumsi IIDN maka model regresi tidak layak menggambarkan hubungan antara variabel prediktor dengan variabel respon

##### 1. Asumsi Identik

Pengujian asumsi residual identik yang artinya tidak terjadi kasus heteroskedastisitas atau variansi residual dari model harus homogeny. Pemeriksaan asumsi residual identik dapat dilakukan

dengan menggunakan uji Glejser. Uji Glejser dilakukan dengan meregresikan absolut dari residual dengan variabel prediktor. Hipotesis yang digunakan untuk uji Glejser terdapat pada persamaan 2.15. Hasil uji Glejser akan ditampilkan pada Tabel 4.9.

**Tabel 4.9** Hasil Uji Glejser

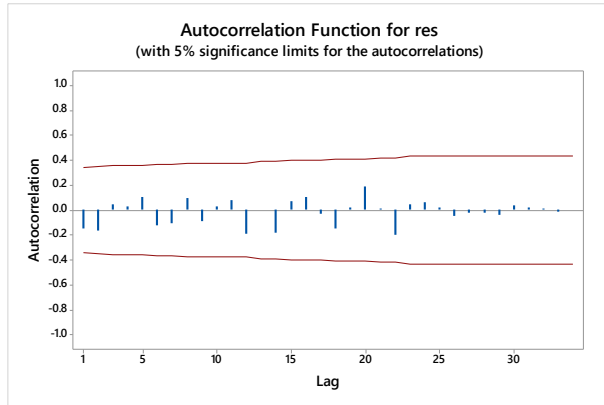
Sumber	df	SS	MS	F <sub>hit</sub>	p-value
Regresi	16	8,872724	0,5545452	0,7607888	0,7062689
Error	18	13,12035	0,7289083		
Total	34	21,99307			

Berdasarkan Tabel 4.9 diketahui bahwa nilai  $p\text{-value}$  sebesar 0,7062689. Dengan taraf signifikansi ( $\alpha$ ) sebesar 5% maka didapatkan keputusan gagal tolak  $H_0$  karena nilai  $p\text{-value} > \alpha$  (0,05) dan nilai  $F_{hitung} < F_{(0,05,16,18)}$  yaitu 0,7607888 < 2,249587, sehingga dapat disimpulkan bahwa tidak ada heteroskedastisitas atau dengan kata lain variansi antar residual sama. Hal ini berarti bahwa asumsi residual identik telah terpenuhi.

## 2. Asumsi Independen

Residual Independen adalah tidak terjadi autokorelasi antar residual. Plot *Autocorrelation Function* (ACF) dapat digunakan untuk mengetahui ada atau tidaknya autokorelasi antar residual. Perhitungan nilai autokorelasi ( $\rho_l$ ) didapatkan dari persamaan 2.17. Batas interval konfidensi yang terdapat pada plot *Autocorrelation Function* (ACF) didapatkan dengan perhitungan menggunakan persamaan 2.18.

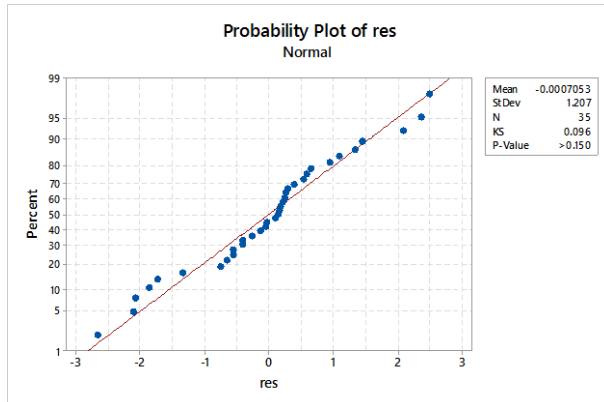
Apabila ada autokorelasi ( $\rho_l$ ) yang keluar dari batas atas maupun batas bawah interval konfidensi maka dapat disimpulkan bahwa terdapat autokorelasi antar residual. Gambar 4.7 menunjukkan bahwa tidak terlihat adanya nilai autokorelasi ( $\rho_l$ ) yang keluar batas interval konfidensi. Hal ini mengindikasikan bahwa asumsi independen pada residual model telah terpenuhi.



**Gambar 4.7** Plot ACF Residual

### 3. Asumsi Distribusi Normal

Pengujian asumsi distribusi normal dilakukan untuk mengetahui apakah residual telah berdistribusi normal atau tidak. Pengujian dapat dilakukan dengan uji *Kolmogorov-Smirnov*, dengan hipotesis yang digunakan terdapat pada persamaan 2.20. Hasil uji *Kolmogorov-Smirnov* ditampilkan pada Gambar 4.8.



**Gambar 4.8** Hasil Uji *Kolmogorov-Smirnov*



Berdasarkan Gambar 4.8 diketahui bahwa nilai *Kolmogorov-Smirnov* sebesar 0,096 dimana nilai ini lebih kecil dibandingkan  $q_{(1-\alpha)}$  yaitu 0,244 dan  $p\text{-value}$  adalah  $> 0,150$  sehingga gagal tolak  $H_0$ . Hal ini menunjukkan bahwa residual model regresi nonparametrik *spline* telah memenuhi asumsi distribusi normal.

#### 4.4 Kriteria Pemilihan Model Terbaik

Salah satu cara untuk menentukan model regresi terbaik ialah menggunakan koefisien determinasi atau  $R^2$ . Untuk mendapatkan nilai  $R^2$  bisa menggunakan rumus yang terdapat pada persamaan 2.7 atau dengan membagi  $SS_{regresi}$  dan  $SS_{total}$  yang diperoleh berdasarkan Tabel 4.7. Dari perhitungan tersebut didapatkan  $R^2$  sebesar 93,1458%. Hal ini menunjukkan bahwa model tersebut mampu menjelaskan keragaman nilai IPM Provinsi Jawa Tengah sebesar 93,1458% sedangkan sisanya dijelaskan oleh variabel lain.

Selain menggunakan  $R^2$ , pemilihan model terbaik juga bisa menggunakan nilai *MSE*. Perhitungan Nilai *MSE* menggunakan rumus yang terdapat pada persamaan 2.8. Sebelumnya nilai *MSE* didapatkan dengan menggunakan metode *k-fold*. Dalam menggunakan perhitungan *k-fold* digunakan  $k=10$ , sehingga didapatkan nilai *MSE* sebesar 6,45564. Nilai *MSE* yang diperoleh termasuk kecil sehingga dikatakan bahwa model yang diperoleh sudah baik.

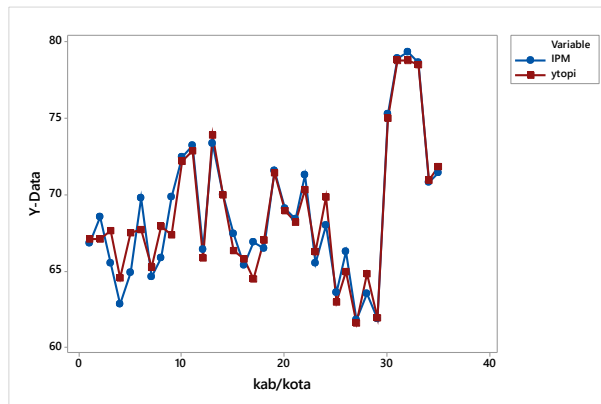
#### 4.5 Pemodelan Regresi Nonparametrik *Spline*

Setelah melakukan pemilihan titik knot optimal menggunakan nilai GCV minimum pada 1 titik knot, 2 titik knot, 3 titik knot dan kombinasi titik knot, didapatkan bahwa kombinasi titik knot merupakan titik knot optimum dengan nilai GCV minimum yaitu 4,80. Sehingga model terbaik regresi nonparametrik *spline* yang akan dipilih adalah model regresi nonparametrik *spline* dengan menggunakan kombinasi titik knot 3,3,2,1,2. Dari model regresi

terbaik yang didapatkan dengan menggunakan kombinasi titik knot (3,3,2,1,2) menghasilkan model sebagai berikut.

$$\begin{aligned}\hat{y} = & 3,699 + 2,561 X_1 - 14,949 (X_1 - 65,73)_+^1 + \\ & 39,517 (X_1 - 66,60)_+^1 - 27,005 (X_1 - 66,89)_+^1 - \\ & 0,057 X_2 + 0,279(X_2 - 321,31)_+^1 - \\ & 0,347(X_2 - 347,01)_+^1 + 0,107(X_2 - 355,58)_+^1 + \\ & 0,005 X_3 - 0,047(X_3 - 3631,86)_+^1 + \\ & 0,043 (X_3 - 3857,63)_+^1 - 12,918 X_4 + \\ & 12,812 (X_4 - 7,20)_+^1 + 0,153 X_5 - \\ & 6,385 (X_5 - 19,09)_+^1 + 6,524(X_5 - 20,29)_+^1\end{aligned}$$

Model diatas sudah dilakukan pengujian signifikansi parameter, secara serentak maupun individu. Didapatkan hasil bahwa semua variabel prediktor yang digunakan dalam pemodelan berpengaruh secara signifikan terhadap variabel respon. Sehingga semua variabel masuk pada pemodelan regresi nonparametrik *spline*. Model diatas juga sudah dilakukan pengujian asumsi residual, dan semua pengujian asumsi residual telah terpenuhi. Dari model diatas didapatkan nilai  $R^2$  sebesar 93,1458% dan nilai  $MSE$  sebesar 6,45564, nilai tersebut sudah cukup menggambarkan bahwa model yang diperoleh sudah baik.



**Gambar 4.9** Nilai  $\hat{Y}_i$  dan  $Y_i$

Untuk melihat apakah model yang didapatkan sudah cukup menggambarkan nilai sebenarnya, dapat dari Gambar 4.9, yang menunjukkan nilai variabel respon ( $Y$ ) dengan nilai yang didapatkan dari model penduga ( $\hat{Y}$ ). Terlihat bahwa nilai  $\hat{Y}_i$  dan  $Y_i$  cenderung berhimpit, hal ini menunjukkan bahwa model yang didapatkan sudah cukup baik untuk menggambarkan nilai  $Y$ .

## LAMPIRAN

**Lampiran 1.** Data Indeks Pembangunan Manusia (IPM) di Jawa Tengah dengan Faktor-Faktor yang Memengaruhi Tahun 2013

<b>Kabupaten</b>	<b>y</b>	<b>x<sub>1</sub></b>	<b>x<sub>2</sub></b>	<b>x<sub>3</sub></b>	<b>x<sub>4</sub></b>	<b>x<sub>5</sub></b>
Kab. Cilacap	66.80	66.31	246.00	784	13.71	56.097880
Kab. Banyumas	68.55	63.95	392.34	1209	19.54	14.237626
Kab. Purbalingga	65.53	73.53	440.20	1131	17.13	8.156952
Kab. Banjarnegara	62.84	73.44	325.50	832	12.74	9.172144
Kab. Kebumen	64.86	71.48	402.72	917	13	8.835316
Kab. Purworejo	69.77	70.87	351.96	682	11.57	8.733568
Kab. Wonosobo	64.57	69.31	336.92	781	17.55	5.327874
Kab. Magelang	65.86	70.23	441.96	1125	13.69	10.814290
Kab. Boyolali	69.81	75.85	437.70	938	18.21	11.168765
Kab. Klaten	72.42	72.68	439.37	1753	13.69	15.217801
Kab. Sukoharjo	73.22	67.89	404.32	1820	9.13	13.760307
Kab. Wonogiri	66.40	71.98	339.61	517	12.52	8.815333
Kab. Karanganyar	73.33	71.04	395.83	1088	6.88	12.857290
Kab. Sragen	69.95	73.68	227.07	921	9	9.684552
Kab. Grobogan	67.43	73.13	402.93	676	14.44	8.934578
Kab. Blora	65.37	75.1	361.04	471	14.8	5.727848
Kab. Rembang	66.84	72.94	333.20	600	14.04	6.579916
Kab. Pati	66.47	70.77	412.45	817	16.1	12.882090
Kab. Kudus	71.58	73.06	451.75	1907	19.43	41.192664
Kab. Jepara	69.11	70.19	356.17	1148	22.79	12.517487
Kab. Demak	68.38	68.11	419.80	1220	13.32	7.950829
Kab. Semarang	71.29	74.14	358.39	1029	12.51	15.748752
Kab. Temanggung	65.52	76.7	330.53	841	8.16	6.915876
Kab. Kendal	67.98	71.86	349.12	925	16.56	14.923437

Kabupaten	y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
Kab. Batang	63.60	70.97	646.87	925	10.83	7.219974
Kab. Pekalongan	66.26	69.37	305.06	1030	15.39	10.014968
Kab. Pemalang	61.81	66.62	438.13	1265	8.7	10.924552
Kab. Tegal	63.50	62.56	412.80	1609	18.45	10.989142
Kab. Brebes	61.87	73.03	416.53	1064	21.6	20.199856
Kota Magelang	75.29	68.37	355.68	6619	8.24	2.911109
Kota Surakarta	78.89	72.1	301.15	11534	7.93	13.599597
Kota Salatiga	79.37	67.96	401.96	3372	12.14	2.282284
Kota Semarang	78.68	67.1	373.57	4402	13.69	61.092826
Kota Pekalongan	70.82	66.22	405.59	6470	8.23	5.201368
Kota Tegal	71.44	70.97	299.13	7070	14.44	3.398773

Keterangan:

y = Indeks Pembangunan Manusia

x<sub>1</sub> = Tingkat Partisipasi Angkatan Kerja

x<sub>2</sub> = Rasio Sekolah-Siswa

x<sub>3</sub> = Kepadatan Penduduk

x<sub>4</sub> = Angka Kesakitan

x<sub>5</sub> = PDRB/1juta

## Lampiran 2. Data Traning *K-fold* 1

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
68.55	63.95	392.34	1209	19.54	14.237626
65.53	73.53	440.20	1131	17.13	8.156952
62.84	73.44	325.50	832	12.74	9.172144
64.86	71.48	402.72	917	13	8.835316
69.77	70.87	351.96	682	11.57	8.733568
64.57	69.31	336.92	781	17.55	5.327874
65.86	70.23	441.96	1125	13.69	10.814290
69.81	75.85	437.70	938	18.21	11.168765
72.42	72.68	439.37	1753	13.69	15.217801
66.40	71.98	339.61	517	12.52	8.815333
73.33	71.04	395.83	1088	6.88	12.857290
69.95	73.68	227.07	921	9	9.684552
67.43	73.13	402.93	676	14.44	8.934578
65.37	75.1	361.04	471	14.8	5.727848
66.84	72.94	333.20	600	14.04	6.579916
66.47	70.77	412.45	817	16.1	12.882090
71.58	73.06	451.75	1907	19.43	41.192664
69.11	70.19	356.17	1148	22.79	12.517487
71.29	74.14	358.39	1029	12.51	15.748752
65.52	76.7	330.53	841	8.16	6.915876
67.98	71.86	349.12	925	16.56	14.923437
63.60	70.97	646.87	925	10.83	7.219974
66.26	69.37	305.06	1030	15.39	10.014968
⋮	⋮	⋮	⋮	⋮	⋮
71.44	70.97	299.13	7070	14.44	3.398773

**Lampiran 3.** Data Testing *K-fold* 1

<b>y</b>	<b>x<sub>1</sub></b>	<b>x<sub>2</sub></b>	<b>x<sub>3</sub></b>	<b>x<sub>4</sub></b>	<b>x<sub>5</sub></b>
66.80	66.31	246.00	784	13.71	56.097880
73.22	67.89	404.32	1820	9.13	13.760307
68.38	68.11	419.80	1220	13.32	7.950829
78.89	72.1	301.15	11534	7.93	13.599597

#### Lampiran 4. Data Traning *K-fold* 2

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
66.80	66.31	246.00	784	13.71	56.097880
65.53	73.53	440.20	1131	17.13	8.156952
62.84	73.44	325.50	832	12.74	9.172144
64.86	71.48	402.72	917	13	8.835316
69.77	70.87	351.96	682	11.57	8.733568
64.57	69.31	336.92	781	17.55	5.327874
65.86	70.23	441.96	1125	13.69	10.814290
69.81	75.85	437.70	938	18.21	11.168765
72.42	72.68	439.37	1753	13.69	15.217801
73.22	67.89	404.32	1820	9.13	13.760307
73.33	71.04	395.83	1088	6.88	12.857290
69.95	73.68	227.07	921	9	9.684552
67.43	73.13	402.93	676	14.44	8.934578
65.37	75.1	361.04	471	14.8	5.727848
66.84	72.94	333.20	600	14.04	6.579916
66.47	70.77	412.45	817	16.1	12.882090
71.58	73.06	451.75	1907	19.43	41.192664
69.11	70.19	356.17	1148	22.79	12.517487
68.38	68.11	419.80	1220	13.32	7.950829
65.52	76.7	330.53	841	8.16	6.915876
67.98	71.86	349.12	925	16.56	14.923437
63.60	70.97	646.87	925	10.83	7.219974
66.26	69.37	305.06	1030	15.39	10.014968
⋮	⋮	⋮	⋮	⋮	⋮
71.44	70.97	299.13	7070	14.44	3.398773



**Lampiran 5.** Data Testing *K-fold* 2

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
68.55	63.95	392.34	1209	19.54	14.237626
66.40	71.98	339.61	517	12.52	8.815333
71.29	74.14	358.39	1029	12.51	15.748752
79.37	67.96	401.96	3372	12.14	2.282284

### Lampiran 6. Data Traning *K-fold* 3

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
66.80	66.31	246.00	784	13.71	56.097880
68.55	63.95	392.34	1209	19.54	14.237626
62.84	73.44	325.50	832	12.74	9.172144
64.86	71.48	402.72	917	13	8.835316
69.77	70.87	351.96	682	11.57	8.733568
64.57	69.31	336.92	781	17.55	5.327874
65.86	70.23	441.96	1125	13.69	10.814290
69.81	75.85	437.70	938	18.21	11.168765
72.42	72.68	439.37	1753	13.69	15.217801
73.22	67.89	404.32	1820	9.13	13.760307
66.40	71.98	339.61	517	12.52	8.815333
69.95	73.68	227.07	921	9	9.684552
67.43	73.13	402.93	676	14.44	8.934578
65.37	75.1	361.04	471	14.8	5.727848
66.84	72.94	333.20	600	14.04	6.579916
66.47	70.77	412.45	817	16.1	12.882090
71.58	73.06	451.75	1907	19.43	41.192664
69.11	70.19	356.17	1148	22.79	12.517487
68.38	68.11	419.80	1220	13.32	7.950829
71.29	74.14	358.39	1029	12.51	15.748752
67.98	71.86	349.12	925	16.56	14.923437
63.60	70.97	646.87	925	10.83	7.219974
66.26	69.37	305.06	1030	15.39	10.014968
⋮	⋮	⋮	⋮	⋮	⋮
71.44	70.97	299.13	7070	14.44	3.398773

**Lampiran 7.** Data Testing *K-fold* 3

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
65.53	73.53	440.20	1131	17.13	8.156952
73.33	71.04	395.83	1088	6.88	12.857290
65.52	76.7	330.53	841	8.16	6.915876
78.68	67.1	373.57	4402	13.69	61.092826

### Lampiran 8. Data Traning *K-fold* 4

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
66.80	66.31	246.00	784	13.71	56.097880
68.55	63.95	392.34	1209	19.54	14.237626
65.53	73.53	440.20	1131	17.13	8.156952
64.86	71.48	402.72	917	13	8.835316
69.77	70.87	351.96	682	11.57	8.733568
64.57	69.31	336.92	781	17.55	5.327874
65.86	70.23	441.96	1125	13.69	10.814290
69.81	75.85	437.70	938	18.21	11.168765
72.42	72.68	439.37	1753	13.69	15.217801
73.22	67.89	404.32	1820	9.13	13.760307
66.40	71.98	339.61	517	12.52	8.815333
73.33	71.04	395.83	1088	6.88	12.857290
67.43	73.13	402.93	676	14.44	8.934578
65.37	75.1	361.04	471	14.8	5.727848
66.84	72.94	333.20	600	14.04	6.579916
66.47	70.77	412.45	817	16.1	12.882090
71.58	73.06	451.75	1907	19.43	41.192664
69.11	70.19	356.17	1148	22.79	12.517487
68.38	68.11	419.80	1220	13.32	7.950829
71.29	74.14	358.39	1029	12.51	15.748752
65.52	76.7	330.53	841	8.16	6.915876
63.60	70.97	646.87	925	10.83	7.219974
66.26	69.37	305.06	1030	15.39	10.014968
⋮	⋮	⋮	⋮	⋮	⋮
71.44	70.97	299.13	7070	14.44	3.398773

**Lampiran 9.** Data Testing *K-fold* 4

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
62.84	73.44	325.50	832	12.74	9.172144
69.95	73.68	227.07	921	9	9.684552
67.98	71.86	349.12	925	16.56	14.923437
70.82	66.22	405.59	6470	8.23	5.201368

### Lampiran 10. Data Traning *K-fold* 5

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
66.80	66.31	246.00	784	13.71	56.097880
68.55	63.95	392.34	1209	19.54	14.237626
65.53	73.53	440.20	1131	17.13	8.156952
62.84	73.44	325.50	832	12.74	9.172144
69.77	70.87	351.96	682	11.57	8.733568
64.57	69.31	336.92	781	17.55	5.327874
65.86	70.23	441.96	1125	13.69	10.814290
69.81	75.85	437.70	938	18.21	11.168765
72.42	72.68	439.37	1753	13.69	15.217801
73.22	67.89	404.32	1820	9.13	13.760307
66.40	71.98	339.61	517	12.52	8.815333
73.33	71.04	395.83	1088	6.88	12.857290
69.95	73.68	227.07	921	9	9.684552
65.37	75.1	361.04	471	14.8	5.727848
66.84	72.94	333.20	600	14.04	6.579916
66.47	70.77	412.45	817	16.1	12.882090
71.58	73.06	451.75	1907	19.43	41.192664
69.11	70.19	356.17	1148	22.79	12.517487
68.38	68.11	419.80	1220	13.32	7.950829
71.29	74.14	358.39	1029	12.51	15.748752
65.52	76.7	330.53	841	8.16	6.915876
67.98	71.86	349.12	925	16.56	14.923437
66.26	69.37	305.06	1030	15.39	10.014968
⋮	⋮	⋮	⋮	⋮	⋮
70.82	66.22	405.59	6470	8.23	5.201368

**Lampiran 11.** Data Testing *K-fold* 5

<b>y</b>	<b>x<sub>1</sub></b>	<b>x<sub>2</sub></b>	<b>x<sub>3</sub></b>	<b>x<sub>4</sub></b>	<b>x<sub>5</sub></b>
64.86	71.48	402.72	917	13	8.835316
67.43	73.13	402.93	676	14.44	8.934578
63.60	70.97	646.87	925	10.83	7.219974
71.44	70.97	299.13	7070	14.44	3.398773

## Lampiran 12. Data Traning *K-fold* 6

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
66.80	66.31	246.00	784	13.71	56.097880
68.55	63.95	392.34	1209	19.54	14.237626
65.53	73.53	440.20	1131	17.13	8.156952
62.84	73.44	325.50	832	12.74	9.172144
64.86	71.48	402.72	917	13	8.835316
64.57	69.31	336.92	781	17.55	5.327874
65.86	70.23	441.96	1125	13.69	10.814290
69.81	75.85	437.70	938	18.21	11.168765
72.42	72.68	439.37	1753	13.69	15.217801
73.22	67.89	404.32	1820	9.13	13.760307
66.40	71.98	339.61	517	12.52	8.815333
73.33	71.04	395.83	1088	6.88	12.857290
69.95	73.68	227.07	921	9	9.684552
67.43	73.13	402.93	676	14.44	8.934578
66.84	72.94	333.20	600	14.04	6.579916
66.47	70.77	412.45	817	16.1	12.882090
71.58	73.06	451.75	1907	19.43	41.192664
69.11	70.19	356.17	1148	22.79	12.517487
68.38	68.11	419.80	1220	13.32	7.950829
71.29	74.14	358.39	1029	12.51	15.748752
65.52	76.7	330.53	841	8.16	6.915876
67.98	71.86	349.12	925	16.56	14.923437
63.60	70.97	646.87	925	10.83	7.219974
⋮	⋮	⋮	⋮	⋮	⋮
71.44	70.97	299.13	7070	14.44	3.398773



**Lampiran 13.** Data Testing *K-fold* 6

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
69.77	70.87	351.96	682	11.57	8.733568
65.37	75.1	361.04	471	14.8	5.727848
66.26	69.37	305.06	1030	15.39	10.014968

# Lampiran 14. Data Traning *K-fold* 7

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
66.80	66.31	246.00	784	13.71	56.097880
68.55	63.95	392.34	1209	19.54	14.237626
65.53	73.53	440.20	1131	17.13	8.156952
62.84	73.44	325.50	832	12.74	9.172144
64.86	71.48	402.72	917	13	8.835316
69.77	70.87	351.96	682	11.57	8.733568
65.86	70.23	441.96	1125	13.69	10.814290
69.81	75.85	437.70	938	18.21	11.168765
72.42	72.68	439.37	1753	13.69	15.217801
73.22	67.89	404.32	1820	9.13	13.760307
66.40	71.98	339.61	517	12.52	8.815333
73.33	71.04	395.83	1088	6.88	12.857290
69.95	73.68	227.07	921	9	9.684552
67.43	73.13	402.93	676	14.44	8.934578
65.37	75.1	361.04	471	14.8	5.727848
66.47	70.77	412.45	817	16.1	12.882090
71.58	73.06	451.75	1907	19.43	41.192664
69.11	70.19	356.17	1148	22.79	12.517487
68.38	68.11	419.80	1220	13.32	7.950829
71.29	74.14	358.39	1029	12.51	15.748752
65.52	76.7	330.53	841	8.16	6.915876
67.98	71.86	349.12	925	16.56	14.923437
63.60	70.97	646.87	925	10.83	7.219974
⋮	⋮	⋮	⋮	⋮	⋮
71.44	70.97	299.13	7070	14.44	3.398773

**Lampiran 15.** Data Testing *K-fold* 7

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
64.57	69.31	336.92	781	17.55	5.327874
66.84	72.94	333.20	600	14.04	6.579916
61.81	66.62	438.13	1265	8.7	10.924552

### Lampiran 16. Data Traning *K-fold* 8

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
66.80	66.31	246.00	784	13.71	56.097880
68.55	63.95	392.34	1209	19.54	14.237626
65.53	73.53	440.20	1131	17.13	8.156952
62.84	73.44	325.50	832	12.74	9.172144
64.86	71.48	402.72	917	13	8.835316
69.77	70.87	351.96	682	11.57	8.733568
64.57	69.31	336.92	781	17.55	5.327874
69.81	75.85	437.70	938	18.21	11.168765
72.42	72.68	439.37	1753	13.69	15.217801
73.22	67.89	404.32	1820	9.13	13.760307
66.40	71.98	339.61	517	12.52	8.815333
73.33	71.04	395.83	1088	6.88	12.857290
69.95	73.68	227.07	921	9	9.684552
67.43	73.13	402.93	676	14.44	8.934578
65.37	75.1	361.04	471	14.8	5.727848
66.84	72.94	333.20	600	14.04	6.579916
71.58	73.06	451.75	1907	19.43	41.192664
69.11	70.19	356.17	1148	22.79	12.517487
68.38	68.11	419.80	1220	13.32	7.950829
71.29	74.14	358.39	1029	12.51	15.748752
65.52	76.7	330.53	841	8.16	6.915876
67.98	71.86	349.12	925	16.56	14.923437
63.60	70.97	646.87	925	10.83	7.219974
⋮	⋮	⋮	⋮	⋮	⋮
71.44	70.97	299.13	7070	14.44	3.398773

**Lampiran 17.** Data Testing *K-fold* 8

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
65.86	70.23	441.96	1125	13.69	10.814290
66.47	70.77	412.45	817	16.1	12.882090
63.50	62.56	412.80	1609	18.45	10.989142

### Lampiran 18. Data Traning *K-fold* 9

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
66.80	66.31	246.00	784	13.71	56.097880
68.55	63.95	392.34	1209	19.54	14.237626
65.53	73.53	440.20	1131	17.13	8.156952
62.84	73.44	325.50	832	12.74	9.172144
64.86	71.48	402.72	917	13	8.835316
69.77	70.87	351.96	682	11.57	8.733568
64.57	69.31	336.92	781	17.55	5.327874
65.86	70.23	441.96	1125	13.69	10.814290
72.42	72.68	439.37	1753	13.69	15.217801
73.22	67.89	404.32	1820	9.13	13.760307
66.40	71.98	339.61	517	12.52	8.815333
73.33	71.04	395.83	1088	6.88	12.857290
69.95	73.68	227.07	921	9	9.684552
67.43	73.13	402.93	676	14.44	8.934578
65.37	75.1	361.04	471	14.8	5.727848
66.84	72.94	333.20	600	14.04	6.579916
66.47	70.77	412.45	817	16.1	12.882090
69.11	70.19	356.17	1148	22.79	12.517487
68.38	68.11	419.80	1220	13.32	7.950829
71.29	74.14	358.39	1029	12.51	15.748752
65.52	76.7	330.53	841	8.16	6.915876
67.98	71.86	349.12	925	16.56	14.923437
63.60	70.97	646.87	925	10.83	7.219974
⋮	⋮	⋮	⋮	⋮	⋮
71.44	70.97	299.13	7070	14.44	3.398773

**Lampiran 19.** Data Testing *K-fold* 9

<b>y</b>	<b>x<sub>1</sub></b>	<b>x<sub>2</sub></b>	<b>x<sub>3</sub></b>	<b>x<sub>4</sub></b>	<b>x<sub>5</sub></b>
69.81	75.85	437.70	938	18.21	11.168765
71.58	73.06	451.75	1907	19.43	41.192664
61.87	73.03	416.53	1064	21.6	20.199856

## Lampiran 20. Data Traning *K-fold* 10

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
66.80	66.31	246.00	784	13.71	56.097880
68.55	63.95	392.34	1209	19.54	14.237626
65.53	73.53	440.20	1131	17.13	8.156952
62.84	73.44	325.50	832	12.74	9.172144
64.86	71.48	402.72	917	13	8.835316
69.77	70.87	351.96	682	11.57	8.733568
64.57	69.31	336.92	781	17.55	5.327874
65.86	70.23	441.96	1125	13.69	10.814290
69.81	75.85	437.70	938	18.21	11.168765
73.22	67.89	404.32	1820	9.13	13.760307
66.40	71.98	339.61	517	12.52	8.815333
73.33	71.04	395.83	1088	6.88	12.857290
69.95	73.68	227.07	921	9	9.684552
67.43	73.13	402.93	676	14.44	8.934578
65.37	75.1	361.04	471	14.8	5.727848
66.84	72.94	333.20	600	14.04	6.579916
66.47	70.77	412.45	817	16.1	12.882090
71.58	73.06	451.75	1907	19.43	41.192664
68.38	68.11	419.80	1220	13.32	7.950829
71.29	74.14	358.39	1029	12.51	15.748752
65.52	76.7	330.53	841	8.16	6.915876
67.98	71.86	349.12	925	16.56	14.923437
63.60	70.97	646.87	925	10.83	7.219974
⋮	⋮	⋮	⋮	⋮	⋮
71.44	70.97	299.13	7070	14.44	3.398773



**Lampiran 21.** Data Testing *K-fold* 10

y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
72.42	72.68	439.37	1753	13.69	15.217801
69.11	70.19	356.17	1148	22.79	12.517487
75.29	68.37	355.68	6619	8.24	2.911109

**Lampiran 22.** Program Pemilihan Titik Knot Optimal dengan Satu Titik Knot Menggunakan *Software R*

```
GCV1=function(para)
{
data=read.table("e://data TA.txt",header=FALSE)
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-para-1
dataA=data[, (para+2):q]
F=matrix(0,nrow=p,ncol=p)
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot1=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot1[j,i]=a[j]
}
}
a1=length(knot1[,1])
knot1=knot1[2:(a1-1),]
aa=rep(1,p)
data1=matrix(ncol=m,nrow=p)
data2=data[,2:q]
a2=nrow(knot1)
GCV=rep(NA,a2)
Rsqr=rep(NA,a2)
for (i in 1:a2)
{
for (j in 1:m)
{
for (k in 1:p)
{
```

```

if (data[k,(j+para+1)]<knot1[i,j]) data1[k,j]=0 else
  data1[k,j]=data[k,(j+para+1)]-knot1[i,j]
}
}
mx=cbind(aa,data2,data1)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====","\n")
cat("Nilai Knot dengan Spline linear 1 knot","\n")
cat("=====","\n")
print(knot1)
cat("=====","\n")
cat("Rsq dengan Spline linear 1 knot","\n")
cat("=====","\n")
print(Rsq)
cat("=====","\n")
cat("HASIL GCV dengan Spline linear 1 knot","\n")

```

```
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print(max(Rsq))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="e://output GCV1.csv")
write.csv(Rsq, file="e://output Rsq1.csv")
write.csv(knot1, file="e://output knot1.csv")
}
```

**Lampiran 23.** Program Pemilihan Titik Knot Optimal dengan Dua Titik Knot Menggunakan *Software R*

```
GCV2=function()
{
data=read.table("e:/data TA.txt", header=FALSE)
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-1
F=matrix(0,nrow=p,ncol=p)
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(data[,i+1]),max(data[,i+1]),length.out=50)
knot[j,i]=a[j]
}
}
z=(nk*(nk-1)/2)
knot2=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot1=rbind(rep(NA,2))
for ( j in 1:(nk-1))
{
for (k in (j+1):nk)
{
xx=cbind(knot[j,i],knot[k,i])
knot1=rbind(knot1,xx)
}
}
knot2=cbind(knot2,knot1)
}
}
```

```

knot2=knot2[2:(z+1),2:(2*m+1)]
aa=rep(1,p)
data2=matrix(ncol=(2*m),nrow=p)
data1=data[,2:q]
a1=length(knot2[,1])
GCV=rep(NA,a1)
Rsqr=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:(2*m))
  {
    if (mod(j,2)==1) b=floor(j/2)+1 else b=j/2
    for (k in 1:p)
    {
      if (data1[k,b]<knot2[i,j]) data2[k,j]=0 else data2[k,j]=data1[k,b]-
        knot2[i,j]
    }
  }
  mx=cbind(aa,data1,data2)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
  SSR=0
  for (r in (1:p))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsqr[i]=(SSR/(SSE+SSR))*100
  MSE=SSE/p
  A=mx%*%C%*%t(mx)
  A1=(F-A)
  A2=(sum(diag(A1))/p)^2

```

```

GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsq)
cat("=====
=====","\n")
cat("Nilai Knot dengan Spline linear 2 knot","\n")
cat("=====
=====","\n")
print (knot2)
cat("=====
=====","\n")
cat("Rsqr dengan Spline linear 2 knot","\n")
cat("=====
=====","\n")
print (Rsqr)
cat("=====
=====","\n")
cat("HASIL GCV dengan Spline linear 2 knot","\n")
cat("=====
=====","\n")
print (GCV)
s1=min(GCV)
cat("=====
=====","\n")
cat("HASIL GCV terkecil dengan Spline linear 2 knot","\n")
cat("=====
=====","\n")
cat(" GCV =",s1,"\n")
write.csv(GCV,file="e://output GCV2.csv")
write.csv(Rsqr,file="e://output Rsqr2.csv")
write.csv(knot2,file="e://output knot2.csv")
}

```

**Lampiran 24.** Program Pemilihan Titik Knot Optimal dengan Tiga Titik Knot Menggunakan *Software R*

```
GCV3=function(para)
{
data=read.table("e://data TA.txt",header=FALSE)
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-para-1
F=matrix(0,nrow=p,ncol=p)
dataA=data[, (para+2):q]
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
knot=knot[2:(nk-1),]
a2=nrow(knot)
z=(a2*(a2-1)*(a2-2)/6)
knot1=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot2=rbind(rep(NA,3))
for ( j in 1:(a2-2))
{
for (k in (j+1):(a2-1))
{
for (g in (k+1):a2)
{
xx=cbind(knot[j,i],knot[k,i],knot[g,i])
knot2=rbind(knot2,xx)
}
}
}
}
}
```



```

}
}
knot1=cbind(knot1,knot2)
}
knot1=knot1[2:(z+1),2:(3*m+1)]
aa=rep(1,p)
data1=matrix(ncol=(3*m),nrow=p)
data2=data[, (para+2):q]
a1=length(knot1[,1])
GCV=rep(NA,a1)
Rsqr=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:ncol(knot1))
  {
    b=ceiling(j/3)
    for (k in 1:p)
    {
      if (data2[k,b]<knot1[i,j]) data1[k,j]=0 else data1[k,j]=data2[k,b]-
        knot1[i,j]
    }
  }
  mx=cbind(aa,data[,2:q],data1)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
  SSR=0
  for (r in (1:p))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsqr[i]=(SSR/(SSE+SSR))*100
  MSE=SSE/p
  A=mx%*%C%*%t(mx)

```

```

A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsqr dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (Rsqr)
r=max(Rsqr)
print (r)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
cat(" GCV =",s1, "\n")
write.csv(GCV,file="e://output GCV3.csv")
write.csv(Rsqr,file="e://output Rsqr3.csv")
write.csv(knot1,file="e://output knot3.csv")
}

```

**Lampiran 25.** Program Pemilihan Titik Knot Optimal dengan Kombinasi Titik Knot Menggunakan *Software R*

```
GCVkom=function(para)
{
data=read.table("e://data TA.txt",header=FALSE)
data=as.matrix(data)
p1=length(data[,1])
q1=length(data[1,])
v=para+2
F=matrix(0,nrow=p1,ncol=p1)
diag(F)=1
x1=read.table("e:/x1.txt")
x2=read.table("e:/x2.txt")
x3=read.table("e:/x3.txt")
x4=read.table("e:/x4.txt")
x5=read.table("e:/x5.txt")
n2=nrow(x1)
a=matrix(nrow=5,ncol=3^5)
m=0
for (i in 1:3)
for (j in 1:3)
for (k in 1:3)
for (l in 1:3)
for (s in 1:3)
{
m=m+1
a[,m]=c(i,j,k,l,s)
}
a=t(a)
GCV=matrix(nrow=nrow(x1),ncol=3^5)
for (i in 1:3^5)
{
for (h in 1:nrow(x1))
{
if (a[i,1]==1)
{
gab=as.matrix(x1[,1])
gen=as.matrix(data[,v])
```

```

aa=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,1]==2)
{
gab=as.matrix(x1[,2:3])
gen=as.matrix(cbind(data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
}
else
{
gab=as.matrix(x1[,4:6])
gen=as.matrix(cbind(data[,v],data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
}
if (a[i,2]==1)
{
gab=as.matrix(x2[,1] )
gen=as.matrix(data[,v+1])
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
}

```

```

}
}
else
if (a[i,2]==2)
{
gab=as.matrix(x2[,2:3] )
gen=as.matrix(cbind(data[, (v+1)], data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x2[,4:6])
gen=as.matrix(cbind(data[, (v+1)], data[, (v+1)], data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,3]==1)
{
gab=as.matrix(x3[,1] )
gen=as.matrix(data[, (v+2)])
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,3]==2)
{

```

```

gab=as.matrix(x3[,2:3] )
gen=as.matrix(cbind(data[, (v+2)], data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x3[,4:6])
gen=as.matrix(cbind(data[, (v+2)], data[, (v+2)], data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
}
if (a[i,4]==1)
{
gab=as.matrix(x4[,1] )
gen=as.matrix(data[, (v+3)])
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
}
else
if (a[i,4]==2)
{
gab=as.matrix(x4[,2:3] )
gen=as.matrix(cbind(data[, (v+3)], data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))

```

```

{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x4[,4:6])
gen=as.matrix(cbind(data[, (v+3)], data[, (v+3)], data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,5]==1)
{
gab=as.matrix(x5[,1] )
gen=as.matrix(data[, (v+4)])
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,5]==2)
{
gab=as.matrix(x5[,2:3] )
gen=as.matrix(cbind(data[, (v+4)], data[, (v+4)]))
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
}
else

```

```

{
gab=as.matrix(x5[,4:6])
gen=as.matrix(cbind(data[, (v+4)], data[, (v+4)], data[, (v+4)]))
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
ma=as.matrix(cbind(aa,bb,cc,dd,ee))
mx=cbind(rep(1,nrow(data)),data[,2:q1],na.omit(ma))
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in 1:nrow(data))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsq=(SSR/(SSE+SSR))*100
MSE=SSE/p1
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p1)^2
GCV[h,i]=MSE/A2
}
if (a[i,1]==1) sp=x1[,1] else
if (a[i,1]==2) sp=x1[,2:3] else
sp=x1[,4:6]
if (a[i,2]==1) spl=x2[,1] else
if (a[i,2]==2) spl=x2[,2:3] else
spl=x2[,4:6]
if (a[i,3]==1) splin=x3[,1] else

```



```

if (a[i,3]==2) splin=x3[,2:3] else
splin=x3[,4:6]
if (a[i,4]==1) spline=x4[,1] else
if (a[i,4]==2) spline=x4[,2:3] else
spline=x4[,4:6]
if (a[i,5]==1) splines=x5[,1] else
if (a[i,5]==2) splines=x5[,2:3] else
splines=x5[,4:6]
kkk=cbind(sp,spl,splin,spline,splines)
cat("=====", "\n")
print(i)
print(kkk)
print(Rsq)
}
write.csv(GCV,file="e://output GCV kombinasi.csv")
write.csv(Rsq,file="e://output Rsq kombinasi.csv")
}

```

## Lampiran 26. Program Kombinasi Titik Knot $3^4$

```
GCVkom=function(data,para)
{
a=matrix(nrow=5,ncol=3^5)
m=0
for (i in 1:3)
for (j in 1:3)
for (k in 1:3)
for (l in 1:3)
for (n in 1:3)

{
m=m+1
a[,m]=c(i,j,k,l,n)
print(a[,m])
}
}
```

**Lampiran 27.** Program Estimasi Parameter dengan Kombinasi Titik Knot 3,3,2,1,2

```

uji=function(alpha,para)
{
data=read.table("e:/data TA.txt")
knot=read.table("e:/knot.txt")
data=as.matrix(data)
knot=as.matrix(knot)
ybar=mean(data[,1])
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m],data[,m],data[,m+1],data[,m+1],data[,m+1],data[,m+2],data[,m+2],data[,m+2],data[,m+3],data[,m+4],data[,m+4])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
for(j in 1:p)
{
if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else data.knot[j,i]=dataA[j,i]-knot[1,i]
}
}
mx=cbind(satu,data[,2],data.knot[,1:3],data[,3],data.knot[,4:6],data[,4],data.knot[,7:8],data[,5],data.knot[,9],data[,6],data.knot[,10:11])
mx=as.matrix(mx)
B=(pinv(t(mx)%*%mx))%*%t(mx)%*%data[,1]
cat("=====","\n")
cat("Estimasi Parameter","\n")
cat("=====","\n")
print (B)
n1=nrow(B)
yhat=mx%*%B
res=data[,1]-yhat
SSE=sum((data[,1]-yhat)^2)

```

```

SSR=sum((yhat-ybar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsq=(SSR/(SSR+SSE))*100

#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan", "\n")
cat("", "\n")
}
else
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh signifikan", "\n")
cat("", "\n")
}

#uji t (uji individu)
thit=rep(NA,n1)
pval=rep(NA,n1)
SE=sqrt(diag(MSE*(pinv(t(mx)%*%mx))))
cat("-----", "\n")
cat("Kesimpulan hasil uji individu", "\n")
cat("-----", "\n")
thit=rep(NA,n1)
pval=rep(NA,n1)
for (i in 1:n1)
{
thit[i]=B[i,1]/SE[i]

```

```

pval[i]=2*(pt(abs(thit[i]),(p-n1),lower.tail=FALSE))
if (pval[i]<=alpha) cat("Tolak Ho yakni prediktor signifikan dengan
pvalue",pval[i],"\n") else cat("Gagal tolak Ho yakni prediktor tidak
signifikan dengan pvalue",pval[i],"\n")
}
thit=as.matrix(thit)
cat("=====", "\n")
cat("nilai t hitung", "\n")
cat("=====", "\n")
print (thit)
cat("Analysis of Variance", "\n")
cat("=====", "\n")
cat("Sumber      df      SS      MS      Fhit", "\n")
cat("Regresi      ",(n1-1), " ",SSR, " ",MSR, " ",Fhit, "\n")
cat("Error        ",p-n1, " ",SSE, " ",MSE, "\n")
cat("Total        ",p-1, " ",SST, "\n")
cat("=====", "\n")
cat("s=",sqrt(MSE), "      Rsq=",Rsq, "\n")
cat("pvalue(F)=",pvalue, "\n")
write.csv(res,file="c:/output uji residual.csv")
write.csv(pval,file="c:/output uji pvalue.csv")
write.csv(mx,file="c:/output uji mx.csv")
write.csv(yhat,file="c:/output uji yhat.csv")
}

```

**Lampiran 28.** Program Uji Glejser untuk Kombinasi Titik Knot  
3,3,2,1,2

```
glejser=function(alpha,para)
{
data=read.table("e:/data TA.txt")
knot=read.table("e:/knot.txt")
res=read.table("e:/res.txt")
data=as.matrix(data)
knot=as.matrix(knot)
res=abs(res)
res=as.matrix(res)
rbar=mean(res)
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m],data[,m],data[,m+1],data[,m+1],data[,m
+1],data[,m+2],data[,m+2],data[,m+3],data[,m+4],data[,m+4])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
for(j in 1:p)
{
if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else data.knot[j,i]=dataA[j,i]-
knot[1,i]
}
}
mx=cbind(satu,data[,2],data.knot[,1:3],data[,3],data.knot[,4:6],data[,4],d
ata.knot[,7:8],data[,5],data.knot[,9],data[,6],data.knot[,10:11])
mx=as.matrix(mx)
B=(pinv(t(mx)%*%mx))%*%t(mx)%*%res
n1=nrow(B)
yhat=mx%*%B
residual=res-yhat
SSE=sum((res-yhat)^2)
SSR=sum((yhat-rbar)^2)
```

```

SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsqr=(SSR/SST)*100

#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
cat("-----","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("-----","\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan atau
terjadi heteroskedastisitas","\n")
cat("","\n")
}
else
{
cat("-----","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("-----","\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh signifikan
atau tidak terjadi heteroskedastisitas","\n")
cat("","\n")
}
cat("Analysis of Variance","\n")
cat("=====","\n")
cat("Sumber      df      SS      MS      Fhit","\n")
cat("Regresi      ",(n1-1)," ",SSR," ",MSR,"",Fhit,"\n")
cat("Error        ",p-n1," ",SSE,"",MSE,"\n")
cat("Total        ",p-1," ",SST,"\n")
cat("=====","\n")
cat("s=",sqrt(MSE),"      Rsqr=",Rsqr,"\n")
cat("pvalue(F)=",pvalue,"\n")
}

```

**Lampiran 29.** Output Nilai GCV dengan Satu Titik Knot pada Lima Variabel Prediktor

No	GCV	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
1	12.5066	62.8486	235.6375	696.7755	7.2047	3.4825
2	14.4784	63.1371	244.2048	922.5510	7.5294	4.6827
3	15.3419	63.4257	252.7722	1148.3265	7.8541	5.8829
4	15.0679	63.7143	261.3395	1374.1020	8.1788	7.0831
5	15.3380	64.0029	269.9068	1599.8776	8.5035	8.2834
6	15.2013	64.2914	278.4742	1825.6531	8.8282	9.4836
7	15.0552	64.5800	287.0415	2051.4286	9.1529	10.6838
8	14.4836	64.8686	295.6089	2277.2041	9.4776	11.8840
9	13.6749	65.1571	304.1762	2502.9796	9.8022	13.0842
10	13.0740	65.4457	312.7435	2728.7551	10.1269	14.2844
⋮	⋮	⋮	⋮	⋮	⋮	⋮
40	20.4668	74.1029	569.7636	9502.0204	19.8678	50.2909
41	20.3470	74.3914	578.3309	9727.7959	20.1924	51.4911
42	20.1894	74.6800	586.8982	9953.5714	20.5171	52.6913
43	19.9573	74.9686	595.4656	10179.3469	20.8418	53.8915
44	19.5414	75.2571	604.0329	10405.1225	21.1665	55.0917
45	18.8027	75.5457	612.6002	10630.8980	21.4912	56.2920
46	18.2462	75.8343	621.1676	10856.6735	21.8159	57.4922
47	18.2222	76.1229	629.7349	11082.4490	22.1406	58.6924
48	18.2222	76.4114	638.3022	11308.2245	22.4653	59.8926



**Lampiran 30.** Output Nilai GCV dengan Dua Titik Knot pada Lima Variabel Prediktor

No	GCV	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
1	13.4634	62.56	227.07018	471	6.88	2.2822837
		62.848571	235.63751	696.77551	7.2046939	3.4824988
2	15.6186	62.56	227.07018	471	6.88	2.2822837
		63.137143	244.20484	922.55102	7.5293878	4.682714
3	16.5862	62.56	227.07018	471	6.88	2.2822837
		63.425714	252.77218	1148.3265	7.8540816	5.8829291
⋮	⋮	⋮	⋮	⋮	⋮	⋮
596	9.6835	66.6	347.01286	3631.8571	11.425714	19.085296
		66.888571	355.58019	3857.6327	11.750408	20.285511
597	10.3445	66.6	347.01286	3631.8571	11.425714	19.085296
		67.177143	364.14753	4083.4082	12.075102	21.485726
598	11.4825	66.6	347.01286	3631.8571	11.425714	19.085296
		67.465714	372.71486	4309.1837	12.399796	22.685941
599	12.2090	66.6	347.01286	3631.8571	11.425714	19.085296
		67.754286	381.2822	4534.9592	12.72449	23.886156
600	13.2311	66.6	347.01286	3631.8571	11.425714	19.085296
		68.042857	389.84953	4760.7347	13.049184	25.086371
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1224	18.2222	76.122857	629.7349	11082.449	22.140612	58.692395
		76.7	646.86957	11534	22.79	61.092826
1225	18.2222	76.411429	638.30223	11308.224	22.465306	59.89261
		76.7	646.86957	11534	22.79	61.092826

**Lampiran 31.** Output Nilai GCV dengan Tiga Titik Knot pada Lima Variabel Prediktor

No	GCV	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
1	16.1837	62.8486	235.6375	696.7755	7.2047	3.4825
		63.1371	244.2048	922.5510	7.5294	4.6827
		63.4257	252.7722	1148.3265	7.8541	5.8829
2	15.0518	62.8486	235.6375	696.7755	7.2047	3.4825
		63.1371	244.2048	922.5510	7.5294	4.6827
		63.7143	261.3395	1374.1020	8.1788	7.0831
3	14.5917	62.8486	235.6375	696.7755	7.2047	3.4825
		63.1371	244.2048	922.5510	7.5294	4.6827
		64.0029	269.9068	1599.8776	8.5035	8.2834
⋮	⋮	⋮	⋮	⋮	⋮	⋮
8931	11.8612	65.7343	321.3109	2954.5306	10.4516	15.4847
		66.3114	338.4455	3406.0816	11.1010	17.8851
		76.4114	638.3022	11308.2245	22.4653	59.8926
8932	6.3632	65.7343	321.3109	2954.5306	10.4516	15.4847
		66.6000	347.0129	3631.8571	11.4257	19.0853
		66.8886	355.5802	3857.6327	11.7504	20.2855
⋮	⋮	⋮	⋮	⋮	⋮	⋮
17295	19.3823	75.5457	612.6002	10630.8980	21.4912	56.2920
		76.1229	629.7349	11082.4490	22.1406	58.6924
		76.4114	638.3022	11308.2245	22.4653	59.8926
17296	18.2396	75.8343	621.1676	10856.6735	21.8159	57.4922
		76.1229	629.7349	11082.4490	22.1406	58.6924
		76.4114	638.3022	11308.2245	22.4653	59.8926

**Lampiran 32.** Output Nilai GCV dengan Kombinasi Titik Knot pada Lima Variabel Prediktor

No	Kombinasi Knot	GCV	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>
1	1 1 1 1 1	12.506623	62.84857	235.6375	696.7755	7.204694	3.482499
2	1 1 1 1 2	18.20177	62.84857	235.6375	696.7755	7.204694	19.0853
							20.28551
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
108	2 1 3 3 3	9.9498915	66.6	235.6375	2954.531	10.45163	15.48465
			66.88857		3631.857	11.42571	19.0853
					3857.633	11.75041	20.28551
109	2 2 1 1 1	13.310411	66.6	347.0129	696.7755	7.204694	3.482499
			66.88857	355.5802			
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
226	3 3 2 1 1	7.4931086	65.73429	321.3109	3631.857	7.204694	3.482499
			66.6	347.0129	3857.633		
			66.88857	355.5802			
227	3 3 2 1 2	4.7998397	65.73429	321.3109	3631.857	7.204694	19.0853
			66.6	347.0129	3857.633		20.28551
			66.88857	355.5802			
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
242	3 3 3 3 2	6.3159735	65.73429	321.3109	2954.531	10.45163	19.0853
			66.6	347.0129	3631.857	11.42571	20.28551
			66.88857	355.5802	3857.633	11.75041	
243	3 3 3 3 3	6.3632355	65.73429	321.3109	2954.531	10.45163	15.48465
			66.6	347.0129	3631.857	11.42571	19.0853
			66.88857	355.5802	3857.633	11.75041	20.28551

### Lampiran 33. Output Estimasi Parameter dan Uji Signifikansi Parameter Model pada Lima Variabel

---

---

#### Estimasi Parameter

---

---

[,1]  
[1,] 3.699289756  
[2,] 2.561337883  
[3,] -14.949053748  
[4,] 39.516958786  
[5,] -27.004968152  
[6,] -0.057214332  
[7,] 0.279315307  
[8,] -0.347285885  
[9,] 0.106653393  
[10,] 0.005136768  
[11,] -0.046931012  
[12,] 0.042841836  
[13,] -12.917558273  
[14,] 12.812272534  
[15,] 0.153420981  
[16,] -6.384684393  
[17,] 6.523688993

---

#### Kesimpulan hasil uji serentak

---

Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan

---

#### Kesimpulan hasil uji individu

---

Tolak Ho yakni prediktor signifikan dengan pvalue 0.0007889328  
Tolak Ho yakni prediktor signifikan dengan pvalue 0.0004008211  
Tolak Ho yakni prediktor signifikan dengan pvalue 0.0003705707  
Tolak Ho yakni prediktor signifikan dengan pvalue 0.0001932133  
Tolak Ho yakni prediktor signifikan dengan pvalue 0.0003341308  
Tolak Ho yakni prediktor signifikan dengan pvalue 0.01831889  
Tolak Ho yakni prediktor signifikan dengan pvalue 0.01235326  
Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.211296

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.5937322  
 Tolak Ho yakni prediktor signifikan dengan pvalue 8.613498e-07  
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.0002753033  
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.0005965105  
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.02277668  
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.02475934  
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.149062  
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.002519944  
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.00217922

=====

nilai t hitung

=====

[,1]  
 [1,] 4.0280734  
 [2,] 4.3328569  
 [3,] -4.3682926  
 [4,] 4.6637459  
 [5,] -4.4150840  
 [6,] -2.5942639  
 [7,] 2.7800502  
 [8,] -1.2961533  
 [9,] 0.5430868  
 [10,] 7.3115653  
 [11,] -4.5027848  
 [12,] 4.1537273  
 [13,] -2.4899918  
 [14,] 2.4496895  
 [15,] 1.5073859  
 [16,] -3.5065261  
 [17,] 3.5719001

Analysis of Variance

=====

Sumber	df	SS	MS	Fhit
Regresi	16	672.7788	42.04868	15.28829
Error	18	49.50693	2.750385	
Total	34	722.2857		

=====

s= 1.658429    Rsq= 93.1458  
 pvalue(F)= 2.448358e-07

### Lampiran 34. Output Uji Glejser

-----  
Kesimpulan hasil uji serentak  
-----

Gagal Tolak  $H_0$  yakni semua prediktor tidak berpengaruh signifikan atau tidak terjadi heteroskedastisitas

Analysis of Variance

```
=====
Sumber      df      SS      MS      Fhit
Regresi     16  8.872724  0.5545452  0.7607888
Error       18 13.12035  0.7289083
Total       34 21.99307
=====
```

$s = 0.8537612$     $Rsq = 40.34327$   
 $pvalue(F) = 0.7062689$

**Lampiran 35.** Hasil *K-fold* 1 dan Nilai MSE

y aktual	y prediksi	MSE
66.80	67.77312	0.94696
73.22	72.35348	0.750852
68.38	67.98215	0.158286
78.89	77.27188	2.618305
MSE total		1.118601

**Lampiran 36.** Hasil *K-fold* 2 dan Nilai MSE

y aktual	y prediksi	MSE
68.55	65.67999	8.236985
66.40	65.73388	0.443716
71.29	69.79967	2.221085
79.37	77.81334	2.4232
MSE total		3.331247



**Lampiran 37.** Hasil *K-fold* 3 dan Nilai MSE

y aktual	y prediksi	MSE
65.53	68.1188	6.701901
73.33	66.21283	50.65414
65.52	66.91621	1.94941
78.68	78.88295	0.04119
MSE total		14.83666

**Lampiran 38.** Hasil *K-fold* 4 dan Nilai MSE

y aktual	y prediksi	MSE
62.84	64.81075	3.883836
69.95	70.85411	0.81741
67.98	71.69674	13.81413
70.82	71.12827	0.095029
MSE total		4.652601

**Lampiran 39.** Hasil *K-fold* 5 dan Nilai MSE

y aktual	y prediksi	MSE
64.86	67.53951	7.179782
67.43	66.35445	1.156813
63.60	60.23629	11.31455
71.44	71.8736	0.188006
MSE total		4.959789

**Lampiran 40.** Hasil *K-fold* 6 dan Nilai MSE

y aktual	y prediksi	MSE
69.77	67.4144	5.548874
65.37	65.98328	0.376111
66.26	64.04789	4.893428
MSE total		3.606138

**Lampiran 41.** Hasil *K-fold* 7 dan Nilai MSE

y aktual	y prediksi	MSE
64.57	64.23311	0.113495
66.84	63.84549	8.967111
61.81	68.99964	51.69089
MSE total		20.25716

**Lampiran 42.** Hasil *K-fold* 8 dan Nilai MSE

y aktual	y prediksi	MSE
65.86	68.29636	5.93584
66.47	67.35342	0.780437
63.50	66.62678	9.776748
MSE total		5.497675

**Lampiran 43.** Hasil *K-fold* 9 dan Nilai MSE

y aktual	y prediksi	MSE
69.81	66.10568	13.722
71.58	71.67998	0.009996
61.87	64.05765	4.785823
MSE total		6.172606

**Lampiran 44.** Hasil *K-fold* 10 dan Nilai MSE

y aktual	y prediksi	MSE
72.42	72.12754	0.085532
69.11	68.81705	0.085821
75.29	74.84232	0.200415
MSE total		0.123923



(Halaman ini sengaja dikosongkan)

## DAFTAR LAMPIRAN

halaman

Lampiran 1	Data Indeks Pembangunan Manusia (IPM) di Jawa Tengah dengan Faktor-Faktor yang Memengaruhi Tahun 2013 .....	47
Lampiran 2	Data Traning <i>K-fold</i> 1 .....	49
Lampiran 3	Data Testing <i>K-fold</i> 1 .....	50
Lampiran 4	Data Traning <i>K-fold</i> 2 .....	51
Lampiran 5	Data Testing <i>K-fold</i> 2 .....	52
Lampiran 6	Data Traning <i>K-fold</i> 3 .....	53
Lampiran 7	Data Testing <i>K-fold</i> 3 .....	54
Lampiran 8	Data Traning <i>K-fold</i> 4 .....	55
Lampiran 9	Data Testing <i>K-fold</i> 4 .....	56
Lampiran 10	Data Traning <i>K-fold</i> 5 .....	57
Lampiran 11	Data Testing <i>K-fold</i> 5 .....	58
Lampiran 12	Data Traning <i>K-fold</i> 6 .....	59
Lampiran 13	Data Testing <i>K-fold</i> 6 .....	60
Lampiran 14	Data Traning <i>K-fold</i> 7 .....	61
Lampiran 15	Data Testing <i>K-fold</i> 7 .....	62
Lampiran 16	Data Traning <i>K-fold</i> 8 .....	63
Lampiran 17	Data Testing <i>K-fold</i> 8 .....	64
Lampiran 18	Data Traning <i>K-fold</i> 9 .....	65
Lampiran 19	Data Testing <i>K-fold</i> 9 .....	66
Lampiran 20	Data Traning <i>K-fold</i> 10 .....	67
Lampiran 21	Data Testing <i>K-fold</i> 10 .....	68
Lampiran 22	Program Pemilihan Titik Knot Optimal dengan Satu Titik Knot Menggunakan <i>Software</i> R .....	69
Lampiran 23	Program Pemilihan Titik Knot Optimal dengan Dua Titik Knot Menggunakan <i>Software</i> R .....	72
Lampiran 24	Program Pemilihan Titik Knot Optimal dengan Tiga Titik Knot Menggunakan <i>Software</i> R .....	75
Lampiran 25	Program Pemilihan Titik Knot Optimal dengan Kombinasi Titik Knot Menggunakan	

	<i>Software R</i> .....	78
Lampiran 26	Program Kombinasi Titik Knot $3^4$ .....	85
Lampiran 27	Program Estimasi Parameter dengan Kombinasi Titik Knot 3,3,2,1,2 .....	86
Lampiran 28	Program Uji Glejser untuk Kombinasi Titik Knot 3,3,2,1,2 .....	89
Lampiran 29	Output Nilai GCV dengan Satu Titik Knot Pada Lima Variabel Prediktor .....	91
Lampiran 30	Output Nilai GCV dengan Dua Titik Knot Pada Lima Variabel Prediktor .....	92
Lampiran 31	Output Nilai GCV dengan Tiga Titik Knot Pada Lima Variabel Prediktor .....	93
Lampiran 32	Output Nilai GCV dengan Kombinasi Titik Knot Pada Lima Variabel Prediktor .....	94
Lampiran 33	Output Estimasi Parameter dan Uji Signifikansi Parameter Model Pada Lima Variabel Prediktor .....	95
Lampiran 34	Output Uji Glejser .....	97
Lampiran 35	Hasil <i>K-fold</i> 1 dan Nilai MSE.....	98
Lampiran 36	Hasil <i>K-fold</i> 2 dan Nilai MSE.....	99
Lampiran 37	Hasil <i>K-fold</i> 3 dan Nilai MSE.....	100
Lampiran 38	Hasil <i>K-fold</i> 4 dan Nilai MSE.....	101
Lampiran 39	Hasil <i>K-fold</i> 5 dan Nilai MSE.....	102
Lampiran 40	Hasil <i>K-fold</i> 6 dan Nilai MSE.....	103
Lampiran 41	Hasil <i>K-fold</i> 7 dan Nilai MSE.....	104
Lampiran 42	Hasil <i>K-fold</i> 8 dan Nilai MSE.....	105
Lampiran 43	Hasil <i>K-fold</i> 9 dan Nilai MSE.....	106
Lampiran 44	Hasil <i>K-fold</i> 10 dan Nilai MSE.....	107
Lampiran 45	Surat Pernyataan Data .....	109

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, dapat disimpulkan beberapa hal sebagai berikut.

1. Pada tahun 2013 Jawa Tengah memiliki nilai IPM sebesar 63,02. Nilai IPM terendah yaitu Kabupaten Pemalang yaitu sebesar 61,81. Nilai IPM tertinggi yaitu Kota Salatiga dengan nilai sebesar 79,37. Berdasarkan kategori IPM yang dikeluarkan oleh PBB, terdapat 11 kabupaten yang masih berada di kategori menengah bawah, dan sisanya berada kategori menengah atas. Belum ada satu kabupaten/kota yang bisa memasuki kategori tinggi, nilai IPM tertinggi di Jawa Tengah pun belum memasuki kategori tinggi.
2. Model regresi nonparametrik *spline* terbaik untuk pemodelan Indeks Pembangunan manusia di Jawa Tengah adalah dengan menggunakan kombinasi knot 3,3,2,1,2. Model ini mempunyai lima variabel yang berpengaruh secara signifikan yaitu tingkat partisipasi angkatan kerja, rasio sekolah-siswa, kepadatan penduduk, angka kesakitan, PDRB/1juta dengan nilai MSE 6,45564 dan nilai koefisien determinasi ( $R^2$ ) 93,1458%. Hal ini menunjukkan bahwa model tersebut mampu menjelaskan keragaman nilai IPM Provinsi Jawa Tengah sebesar 93,1458% sedangkan sisanya dijelaskan oleh variabel lain. Model regresi nonparametrik *spline* yang diperoleh adalah sebagai berikut.

$$\begin{aligned}\hat{y} = & 3,699 + 2,561 X_1 - 14,949 (X_1 - 65,73)_+^1 + \\ & 39,517 (X_1 - 66,60)_+^1 - 27,005 (X_1 - 66,89)_+^1 - \\ & 0,057 X_2 + 0,279(X_2 - 321,31)_+^1 - \\ & 0,347(X_2 - 347,01)_+^1 + 0,107(X_2 - 355,58)_+^1 + \\ & 0,005 X_3 - 0,047(X_3 - 3631,86)_+^1 + \\ & 0,043 (X_3 - 3857,63)_+^1 - 12,918 X_4 + \\ & 12,812 (X_4 - 7,20)_+^1 + 0,153 X_5 - \\ & 6,385 (X_5 - 19,09)_+^1 + 6,524(X_5 - 20,29)_+^1\end{aligned}$$

## 5.2 Saran

Beberapa saran yang dapat diberikan oleh penulis yaitu sebagai berikut.

1. Bagi penelitian selanjutnya, sebaiknya menambah jumlah variabel yang digunakan dan diduga berpengaruh terhadap Indeks Pembangunan Manusia (IPM) di Jawa Tengah, sehingga diharapkan akan mendapatkan model yang lebih sesuai.
2. Bagi pemerintah, sebaiknya lebih memerhatikan ketiga dimensi yang terdapat pada Indeks Pembangunan Manusia (IPM). Membuat program-program yang menunjang nilai dari ketiga dimensi yang sekiranya bisa meningkatkan angka Indeks Pembangunan Manusia (IPM) terutama di provinsi Jawa Tengah.

## DAFTAR PUSTAKA

- Awal, C. P. (2014). *Pendekatan Regresi Nonparametrik Spline untuk Memodelkan Indeks Pembangunan Manusia (IPM) di Papua*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Badan Pusat Statistika . (2015). *Jawa Tengah dalam Angka 201*. Jawa Tengah: Badan Pusat Statistika.
- Badan Pusat Statistika. (2007). *Indeks Pembangunan Manusia 2006-2007*. Jakarta: Badan Pusat Statistik Indonesia.
- Badan Pusat Statistika. (2011). *Indeks Pembangunan Manusia 2010-2011*. Jakarta: Badan Pusat Statistik Indonesia.
- Badan Pusat Statistika. (2015). *Statistika Daerah Jawa Tengah*. Jawa Tengah : BPS Provinsi Jawa Tengah.
- Budiantara , I. (2000). Metode U, GLM, CV, dan GCV dalam Regresi Nonparametrik. *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)*, 6, 41-45.
- Budiantara, I. (2005). Model Keluarga Spline Polinomial Truncated Dalam Regresi Semiparametrik. *Makalah Seminar Nasional Matematika*, Jurusan Matematika Universitas Diponegoro, Semarang.
- Budiantara, I. (2006). Model Spline dengan Knot Optimal. *Jurnal Ilmu Dasar, FMIPA, Universitas Jember*, 7, 77-85.
- Budiantara, I. (2009). *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Surabaya: ITS Press.
- Drapper, N. R., & Smith, H. (1992). *Analisis Regresi Terapan*. Jakarta: Gramedia Pustaka Utama.
- Eubank , R. (1988). *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker Inc.
- Eubank , R. (1999). *Nonparametric Regression and Spline Smoothing*. New York: Marcel Dekker Inc.
- Gujarati, D. N. (2003). *Basic Econometrics 4th edition*. New York: McGraw-Hill Inc.

- Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). *Applied Linear Regression Models 4th Edition*. New York: The McGraw-Hill Company, Inc.
- Melliana, A. (2013). *Analisis Statistika Faktor yang Mempengaruhi IPM di kabupaten/kota provinsi Jawa Timur dengan Menggunakan regresi panel*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Preston, Samuel, Heuveline, P., & Guillot, M. (2004). *Mearing and Modelling Population Processes*. USA: Blackwell.
- Retno, A. T. (2014). *Faktor-Faktor yang mempengaruhi Indeks Pembangunan Manusia di Jawa Timur dengan Pendekatan Regresi Semiparametrik Spline*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Sugiyono. (2010). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: Alfabeta.
- United Nations Development Programme (UNDP). (1990). *Global Human Development Report*. Human Resources Department.
- Wahba, G. (1990). *Spline Models for Observation Data*. SIAM Pennsylvania.

## BIODATA PENULIS



Ni Putu Dera Yanthi lahir di Bangli, 17 Mei 1994. Putri pertama dari I Nyoman Dena Asa dan Ni Ketut Ranten ini memiliki hobi jalan-jalan, teateran, dan makan. Penulis telah menempuh pendidikan formal di SD Negeri 3 Kesiman, SMPN 9 Denpasar, dan SMA Negeri 1 Denpasar. Penulis melanjutkan ke jenjang perguruan tinggi yaitu di jurusan Statistika ITS melalui jalur SNMPTN undangan. Semasa perkuliahan, penulis mengikuti organisasi baik di dalam atau diluar perkuliahan yaitu Sekertaris Departemen Seni TPKH-ITS periode kepengurusan 2013/2014 dan anggota Komunitas Srikandi Project. Selama perkuliahan, penulis aktif dalam beberapa kepanitian didalam atau diluar kampus, baik menjadi panitia inti maupun anggota sie. Penulis pernah menjadi sutradara dalam 2 pementasan operet yang diadakan di Surabaya. Untuk berdiskusi lebih lanjut mengenai tugas akhir, hubungi penulis melalui:  
Email : [derayanthi@gmail.com](mailto:derayanthi@gmail.com)